DISTINCT ROLES FOR DNMT1 IN MODUATING THE CANCER EPIGENOME

By

Yung-Sheng Melody Tsui

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
September 2015

ABSTRACT

DNA methylation is an epigenetic modification involved in key physiological processes, including X chromosome inactivation, imprinting, and transcriptional regulation in higher eukaryotes. The methylation patterns are formed early in embryonic development and copied during somatic cell division by a family of enzymes known as DNA methyltransferases (DNMTs). Aberrant DNMT function contributes to hypermethylation of promoter CpG islands in tumor suppressor genes, and hypomethylation of repetitive sequences in the genome. These epigenetic changes play a significant role in various diseases.

While there are three major DNMT isoforms, DNMT1, DNMT3a and DNMT3b, the degree of redundancy of function at specific locations in the genome is not well understood. We have performed genome wide DNA methylation and gene expression analysis using MBD-Seq and RNA-Seq respectively on genomic DNA obtained from a set of isogenic cell lines in which either DNMT1, DNMT3b, both or neither were genetically disrupted. Data from bisulfite sequencing was used to confirm methylation differences between these isogenic lines. These data revealed that while disruption of both DNMT1 and DNMT3b was required to prevent maintenance of methylation at most genomic sites, maintenance of methylation at some genomic regions could be abrogated by disruption of just DNMT1 or DNMT3b alone. We found that while DNMT3b was required for maintaining methylation particularly at many coding exons, DNMT1 was required for maintenance of methylation largely at putative promoter regions. Additionally, DNMT1's methylating ability was independent of clonal selection, and

DNMT1 mediated methylation led to gene repression in a subset of genomic regions. Finally, we discovered that DNMT1 could also control gene expression in the absence of *cis* DNA methylation alterations, providing evidence that it could act as a direct transcriptional repressor for a subset of genes.

Identifying DNMT1's methylation dependent and independent roles in gene silencing is an important aspect of understanding cancer epigenetics. Our data help to clarify DNMT1's role in in this regard.

Advisor/Reader: Dr. William G. Nelson

Co-Advisor/Reader: Dr. Srinivasan Yegnasubramanian

Thesis Committee Members: Dr. Stephen B. Baylin

Dr. William B. Isaacs

ACKNOWLEDGEMENTS

These past seven years have been both a blessing and a challenge. I am not the same person I was when I first began this journey, and I owe incredible debts of gratitude to everyone who has walked me through these years and made me who I am today.

First and foremost, I want to thank my advisors, William Nelson and Vasan Yegnasubramanian, for providing me with the mentorship and freedom to grow as a scientist. I am deeply grateful for all your support. Within the lab, I would like to acknowledge Michael Haffner, David Esopi, and Sunil Gangadharan for all the technical assistance they have given me. And to my past and present lab mates- Traci Speed, Grace Cornblatt, Nick Wyhs, Dave Walker, Debika Shinohara, Jianyong Liu, Chris Weier, and Nicki Castagna, you have all contributed to making this experience rewarding and enjoyable.

I must thank my best friends who have stuck by me throughout the years. Ellen and Kankan, I truly could not have done this without you girls- your friendship means the world to me.

I am grateful to my parents, Joseph and Julia Tsui, for always knowing and wanting what is best for me, even when I don't yet see it for myself. Thank you for everything you have sacrificed for me, and for teaching me the values of kindness, patience, and perseverance.

Last but not least, I would like to thank my husband, Aaron Priest, for giving me the support I needed to power through to the end. Thank you for being my best friend and loving husband. I can't wait to see what life holds for us next.

DEDICATION

To my parents, Joseph and Julia Tsui, for all the sacrifices they have made so I can be here today, for their unconditional love and support, and most of all, for raising me to fear God and place Him first in all aspects of my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

1.1 A Brief Overview of Epigenetics

The human genome provides information in two forms: genetic and epigenetic. While the genetic information provides the blueprints for making all living organisms, the epigenetic information provides the instructions for carrying it out. Epigenetics is the manual on how, when, and where the genetic information should be used.

In short, epigenetics is the study of heritable changes in gene expression that occur without change in DNA sequence. Epigenetic signals include DNA methylation, histone modifications such as acetylation, phosphorylation, and methylation, and incorporation of histone variants and non-coding RNAs. All these signals cooperate to modulate chromatin structure and determine the transcriptional activity of the genome. Epigenetics impacts many areas of scientific research, including somatic gene therapy [1], cloning [2], cancer biology [3], genomic imprinting [4], and developmental abnormalities [5, 6].

A focal point in research for the past two decades have been the three main, inter-related types of epigenetic inheritance: DNA methylation, genomic imprinting, and histone modification. Events such as X-chromosome inactivation and genomic imprinting were found to depend on molecular events such as methylation and histone hypoacetylation [7-13]. The inheritance of these epigenetic states also depends on DNA methylation during cell division, another direct epigenetic mechanism for maintaining gene repression. The

epigenetic templates that control gene expression are transmitted to daughter cells independently of the DNA sequences. However, these metastable patterns can sometimes become abnormal both in fetal development and the normal aging process, contributing to the common cancer risk. Because of this, we aim to study as much about these processes as we can, in the hope of defeating this 'Emperor of all maladies [14].'

1.2 DNA Methylation

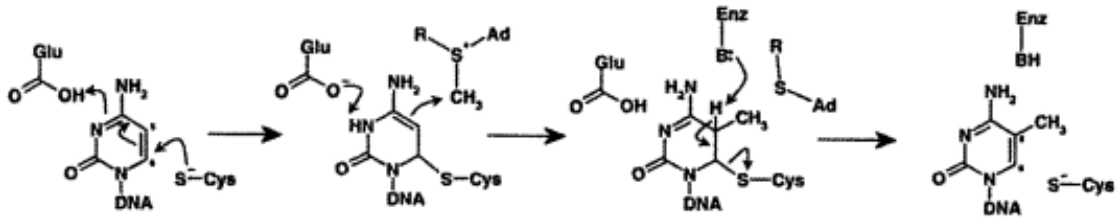
The major DNA modification involved in encoding epigenetic information in mammalian cells is DNA methylation, a reaction that involves the covalent addition of a methyl group to the 5-postion of cytosine. First discovered in calf thymus DNA in 1948 [15], DNA methylation has proven to be a key epigenetic process involved in the control of gene expression, regulation of parental imprinting [4], and stabilization of X-chromosome inactivation [7-13]. It is also implicated in the development of the immune system [16], cellular reprogramming [17-19], and in brain function and behavior [20, 21]. Aberrant DNA methylation patterns are associated with several diseases, among them the initiation and progression of cancer [22-24].

In mammals, DNA methylation occurs primarily in CG dinucleotides, though it occasionally is found in non-CG sites. There are approximately 56 million CG sites in the human genome, about 60-80% of which are methylated [25, 26]. Methylation patterns vary with cell types, with the largest deviations seen in embryonic stem cells [25]. Much of the CG sites have been depleted from the human genome in the course of evolution

due to the mutagenic property of the 5-methyl-cytosine base. However, there remain regions of CpG islands [27] where the ratio of observed to expected CpG is greater than 0.65. These CpG islands are found in promoter regions of about 70% of human genes [28] and DNA methylation at these promoter regions leads to repression of gene activity [25, 26, 29]. In contrast, active genes usually show hypomethylation around the transcriptional start site and high levels of methylation in the gene body [30]. In addition, DNA methylation levels change slightly at splicing sites, with exons showing higher rates of methylation than introns [25, 30], suggesting a role for DNA methylation in splicing as well.

DNA methylation is catalyzed by a group of proteins known as DNA methyltransferases (DNMTs). These enzymes catalyze the transfer of the methyl group from S-adenosyl-L-methionine (SAM) to the C5 position of cytosine residues, creating 5-methylcytosine. Figure 1.1 depicts the reaction mechanism. The methyl addition does not interfere with Watson Crick base pairing because it occurs in the major groove of DNA, where DNA-interacting proteins can also recognize it efficiently.

Figure 1.1: ***Mechanism of Action of DNA Methyltransferase.*** The reaction begins with the cytosine base flipped out of the DNA and inserted into the binding pocket of DNMT. The catalytic cysteine thiolate forms a transition state intermediate with the C6 of cytosine. The resulting 4-5 enamine attacks the methyl group provided by S-Adenosylmethionine (SAM). After addition of the methyl group to C5 of cytosine, a proton is abstracted from the 5 position leading to reformation of the 5,6 double bond and beta elimination of the enzyme.



1.3 DNA Methyltransferases: Form and Function

Global DNA methylation patterns are established by a group of proteins known as DNA methyltransferases, or DNMTs. The mammalian DNMT family contains four active members: DNMT1, DNMT3a, DNMT3b, and an interacting protein DNMT3L. The initial methylation pattern is set by the DNMT3 family [31, 32], whose de novo activity is indicated by their ability to target both hemimethylated and unmethylated DNA [33, 34]. After each round of DNA replication, the parental strand carries the original methylation marks while the daughter strand is unmethylated. This hemimethylated DNA is then targeted by a maintenance methyltransferase (DNMT1) that copies the methylation marks from the parental strand to the daughter strand. While DNMT1 has significant catalytic preference for methylated hemimethylated DNA, its catalytic activity at unmethylated DNA is only modestly decreased compared to that of DNMT3A and DNMT3b. Thus, DNMT1 may participate in both de novo and maintenance methylation. A fifth member, DNMT2, lacks an N-terminal regulatory domain and has the potential to methylate tRNA instead of DNA [35].

Mammalian DNMTs are comprised of two parts: an N-terminal regulatory domain and a C-terminal catalytic domain. The N-terminal domain guides nuclear localization of the enzymes and mediates their interaction with other proteins, DNA, and chromatin. For instance, the N-terminal regulatory region of DNMT1 contains a proliferating cell nuclear antigen-binding domain (PBD), a nuclear localization signal (NLS), a cysteine-rich

ATRX zinc finger DNA-binding motif and a polybromo homology domain (PHD) targeting DNMT1 to the replication foci. Meanwhile, the smaller C-terminal part is the catalytic center of the enzyme (Figure 1.2) [36].

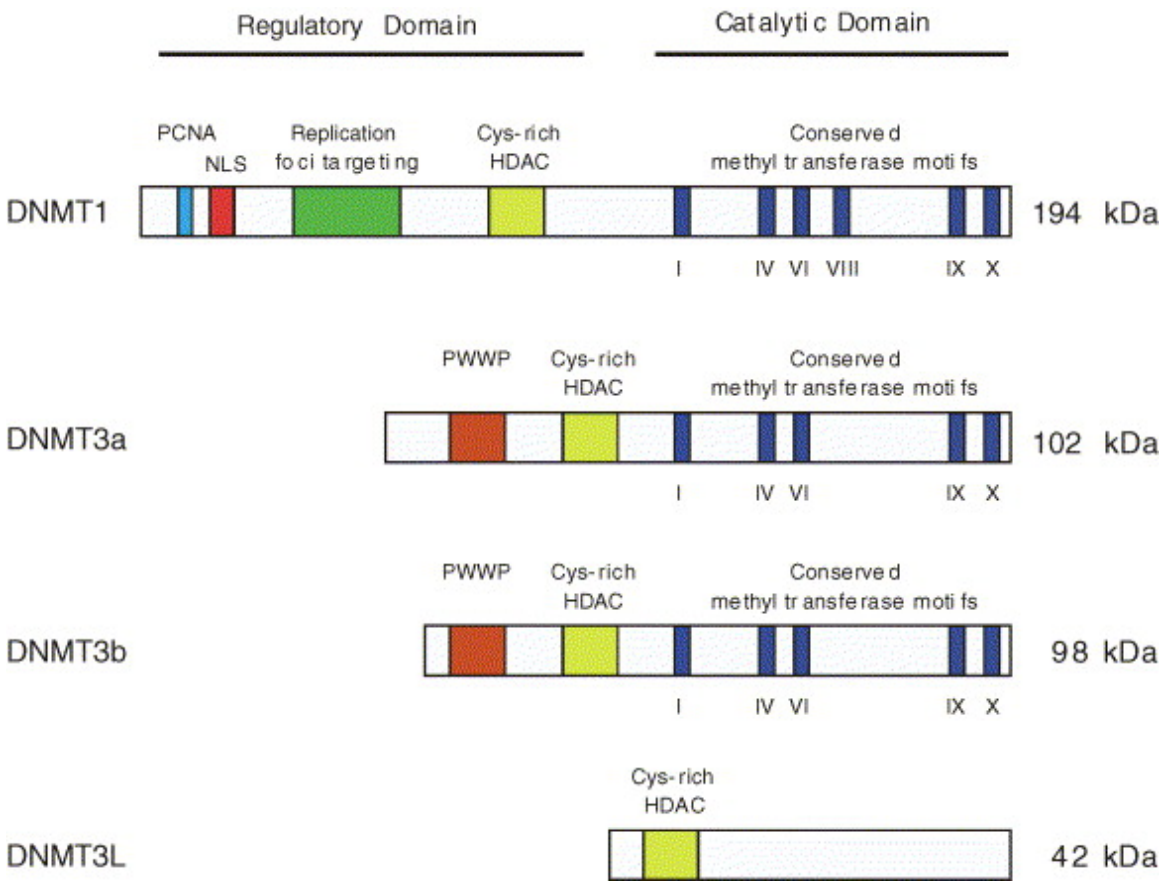
DNMT1 is the most abundant methyltransferase [37] involved in maintenance methylation [38, 39]. DNMT1 shows a preference for hemimethylated DNA and is localized at the DNA replication fork during S phase [40]. It is a highly processive enzyme, able to catalyze multiple reactions before dissociating from the DNA template [41], a property that fits well with its function at the replication fork. It is also the most abundant methyltransferase in somatic cells [42]. DNMT1 is essential to development, as its disruption results in extensive demethylation of the genome and embryonic lethality in mice [5]. As previously discussed, mutations in DNMT1 or its loss leads to alterations imprinting [4] and in X-chromosome inactivation [7-13], suggesting that DNMT1 is required for the maintenance of imprinting signals and stable X-chromosome inactivation. Disruption of DNMT1 in human colorectal carcinoma cells results in loss of proliferation and severe mitotic defects, followed by cell death [43]. Furthermore, DNMT1 loss of function is directly linked to tumorigenesis, as demonstrated in the finding that mice with depleted levels of DNMT1 are susceptible to tumors and display chromosomal instability [44, 45]. These represent just a small sampling of the gene targeting studies that have been done on this protein. Even so, one can already see that DNMT1 plays a crucial role in mammalian development, cell proliferation, and survival.

The Dnmt3 protein family comprises of 3 members: Dnmt3a, Dnmt3b, and Dnmt3L. DNMT3a and DNMT3b are known as de novo methyltransferases because they have an equal preference for hemimethylated and unmethylated DNA substrates [33, 34]. They are responsible for establishing methylation patterns early in mammalian development and in germ cells, and are essential to embryonic development in mice [46]. DNMT3b knockout mouse embryos die in utero and DNMT3a knockout mice die shortly after birth, revealing that disruption of any active DNMT is lethal. Mutations in human DNMT3b also causes ICF syndrome [47], a condition characterized by immunodeficiency, facial defects, and chromosomal abnormalities.

Possessing homology to DNMT3a and DNMT3b, DNM3L serves as structural support for its de novo siblings. Although DNMT3L lacks its own catalytic activity, it is able to stimulate the activity of DNMT3a/3b in vivo by increasing their ability to bind to the methyl group donor, S-adenosyl-L-methionine (SAM) [48]. DNMT3L is thought to be expressed largely during gametogenesis and during various embryonic stages [49].

Cooperation among the DNMTs is required for maintaining methylation in certain genomic regions. Historically, DNMT1 has always been known as the maintenance methyltransferase, while DNMT3a and DNMT3b act as de novo enzymes. However, recent evidence indicates that DNMT1 may also be required for de novo methylation of genomic DNA [50, 51], and that the DNMT3 family members contribute to maintaining methylation during replication [52, 53].

Figure 1.2: ***Schematic of Mammalian DNMT Family Members*.*** Functional domains in the N-terminal part of the proteins are shown and the conserved C5 DNA MTase motifs in the C-terminal domains are labeled.



* Figure adapted from: R Villa, F De Santis, A Gutierrez, S Minucci, P.G Pelicci, L Di Croce, Epigenetic gene silencing in acute promyelocytic leukemia, Biochemical Pharmacology, Volume 68, Issue 6, 15 September 2004, Pages 1247-1254, ISSN 0006-2952

1.4 Aberrant methylation in cancer

In 1993, Holliday et al first indicated that DNA methylation could be an epigenetic mechanism for carcinogenesis [54]. Today, it is widely accepted that abnormal silencing of genes in cancer cells is mediated by aberrant methylation patterns. Specifically, methylation of CpG islands in gene promoter regions, coupled with gene body hypomethylation is associated with transcriptional silencing in tumor suppressor genes [23, 55, 56]. Promoter regions, normally hypomethylated, allow for an open chromatin configuration that permits easy access by transcription factors and RNA polymerase II to bind to the promoter and facilitate gene expression. In contrast, the hypermethylated promoter regions in cancer cells blocks transcription factor binding. Furthermore, loss of CpG methylation within the gene body in cancer cells causes loss of imprinting of genes and chromosomal instability, leading to tumorigenesis. Many genes modified by promoter hypermethylation have tumor suppressor function. Examples include the VHL gene in renal cancer, the cell cycle control gene p16 in many cancers, and the mismatch repair gene MLH1 in colorectal cancer [57, 58].

In cancer cells, DNMT1 seems to be responsible for most of the DNA methylating capacity [59, 60]. In mouse models, DNMT1 is critical for development of intestinal polyposis, and disruption of the DNMT1 gene drastically reduces the incidence of tumors [61, 62]. Recent studies also suggest an interaction between DNMT1 and DNMT3b in colon cancer cells [59, 60], and that DNMTs may contribute to transcriptionally repressive chromatin by mechanisms other than methylation by forming complexes in the

promoter regions with groups of proteins that act to prevent the transcription of genes [63-69].

In prostate cancer, transformation of normal prostatic epithelium to invasive adenocarcinoma can begin at early precursor states called PIA and PIN [70-72]. PIA and PIN lesions begin to exhibit hypermethylation in CpG island sequences of GSTP1, a protein that encodes a single class of glutathione S-transferases that attenuates the development of cancer upon exposure to carcinogens [73]. Hypermethylation increases as prostate cancer progresses, suggesting that chronic or recurrent inflammation may play a role in de novo acquisition of DNA methylation patterns.

Because of the strong association between aberrant methylation and carcinogenesis, much research has focused on elucidating the factors that control DNA methylation. Understanding how DNMTs set up and maintain methylation patterns is a crucial part of the process. While the basic function of DNMTs has been revealed over time, the mechanism by which their activity is modulated is still unclear. For instance, do individual DNMTs act alone or cooperatively? Which parts of the genome do they target? Can DNMTs mediate gene expression independent of methylation? This dissertation explores the roles of DNMT1 and DNMT3b in maintaining DNA methylation patterns, and explores evidence suggesting that DNMT1 can mediate gene repression independent of its DNA methylation activity.

II. DNMT1 REQUIRING DNA METHYLATION

2.1 Introduction

Three DNMTs are known to participate in generating and maintaining DNA methylation patterns: DNMT1, DNMT3a, and DNMT3b. Aberrant expression of these DNMTs disrupts normal methylation patterns and is closely associated with many forms of disease. For instance, hypermethylation at promoter CpG islands can lead to transcriptional silencing of key cellular pathways. In prostate cancer, methylation of the CpG-rich promoter region of GSTP1 is very common, and is one among many other hypermethylated genes [70-73]. These aberrant methylation patterns have also been linked to almost all other known cancer types including colorectal, breast, and lung cancers [74], suggesting that epigenetic silencing tumor suppressor and caretaker genes is an important contributor to the development of cancer. Meanwhile, inhibition of DNMTs correlates with reduction in tumorigenicity and increased expression of tumor suppressor genes [75]. Therefore, an in depth knowledge of how these methylation patterns are established and maintained is a vital step that can lead to disease treatment and diagnostic strategies.

Various methods for genome-wide analysis of DNA methylation have been developed and utilized. Popular tools for differentiating methylated from unmethylated DNA include the use of sodium bisulfite modification, methylation sensitive and specific restriction enzymes (HpaII and McrBcI, respectively), an anti-5mC antibody, or methyl-

binding domain (MBD) polypeptides [76]. These approaches have been successfully used in our lab and others to generate methylation profiles of various normal and cancer cell lines [77-79]. In particular, affinity based strategies like the use of MBD polypeptides are attractive because they are not limited to specific sequences, generate a positive signal for methylation, and are highly effective in separating unmethylated from methylated DNA [80, 81]. MBD polypeptides can also recognize 5-methyl-C in double stranded DNA, which, when combined with next generation sequencing [106], allowed us to generate global methylome profiles of a cancer cell line model with targeted disruption of DNMT1 and DNMT3b.

While the functions of DNMT1, DNMT3a, and DNMT3b have been the subject of muc research, there is still much unknown about their roles in mediating methylation of distinct genome compartments. In order to elucidate the role of DNMTs in genome methylation, we sought to determine how such methylation patterns can be affected by targeted genetic disruption of certain DNMTs. Using MBD-seq, we took isolated DNA from a set of HCT116 isogenic cell lines [59, 60], and subjected them to MBD based enrichment followed by next-generation sequencing. The following cell lines were used: HCT116-*WT*, HCT116-*DNMT1-/-,* HCT116-*DNMT3b-/-,* and HCT116-*DNMT1-/- DNMT3b-/-* (DKO). The use of MBD-seq allowed us to develop genome-wide maps of methylated regions in each of these cell models and to identify regions of altered methylation in each of the DNMT knockout cells compared to the HCT116-WT parental cells. By comparing the methylation profile of the four HCT116 cell lines, we found regions of differential methylation mediated by one or more DNMTs. These methylomes

provide the foundation for which we examine the role of DNMT1 in the genome. In the following chapter, we will discuss the requirement of DNMT1 and DNMT3b in maintaining methylation at various genomic regions, and examine if DNMT1 dependent methylation can lead directly to gene repression.

2.2 MATERIALS AND METHODS

2.2.1 Cell Culture

Experiments were performed using a colorectal carcinoma line, HCT116, in which DNMT1, DNMT3b or both or neither were genetically disrupted. All cell lines were grown in McCoy's medium with 10% FBS at 37 degrees.

2.2.2 siRNA Treatment

Cells were grown in 6 well plates and treated with 25nM of a siRNA pool (4) targeted to DNMT1 (Qiagen) every 24hrs for 72hrs. Cells were harvested 24hrs after the last transfection. McCoy's serum free media was added to cells immediately prior to transfection. A transfection complex composed of Lipofectamine 2000, siDNMT1, and Optimem made according to the manufacturer's protocol were allowed to sit at room temperature for 20 minutes before being applied to cells in serum free media. Five hours later, transfection complexes were removed and the medium was replaced with McCoy's medium with 10% FBS.

2.2.3 Decitabine Treatment

Decitabine, or 5-aza-2-deoxycytidine (Sigma) was dissolved in DMSO immediately prior to use. Five concentrations (50nM, 100nM, 250nM, 500nM, and 1μM) of Decitabine

were added to cells seeded 24hrs prior in McCoy's medium with 10% FBS. Cells were harvested for analysis 24hrs post addition of Decitabine.

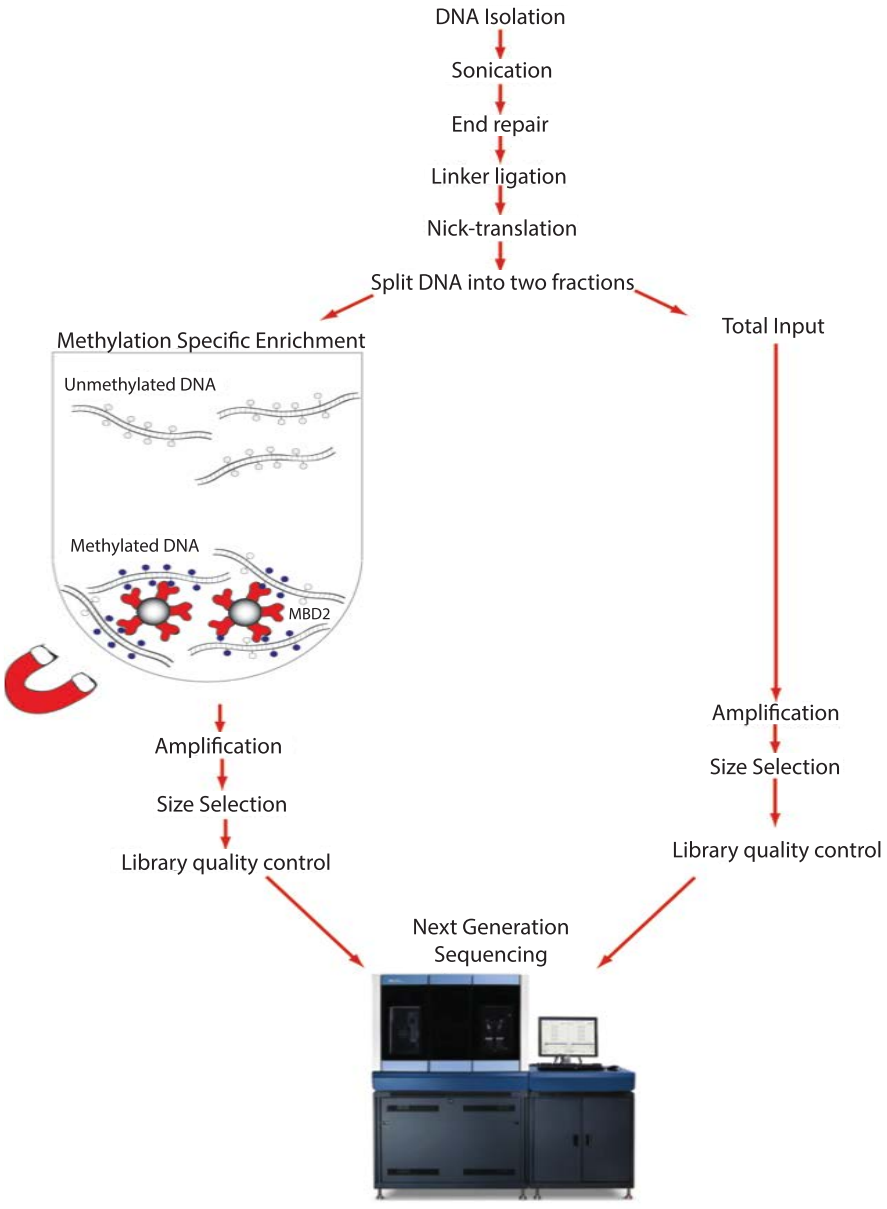
2.2.4 MBD-Seq

A schematic of the protocol is diagrammed in Figure 2.1. Two micrograms of genomic DNA from HCT116-WT, HCT116-*DNMT1-/-,* HCT116-*DNMT3b-/-* and HCT116-DKO were sonicated to a modal size of ~150–250 bp, and end-repaired using the NEBNext SOLiD DNA library preparation kit end-repair module following the manufacturer's protocol (New England Biolabs). After column-purification (using the Qiagen PCR purification kit), SOLiD P1 and P2 adapters lacking 5′ phospate groups (Life Technologies) were ligated using the NEBNext adaptor ligation module, column-purified, and treated with Platinum Taq polymerase for nick translation. The resulting library was divided into two fractions, a total input fraction, and an enriched methylated fraction. The methylated DNA was separated from unmethylated DNA in the enriched fraction by affinity enrichment of methylated DNA fragments using 6xHis-MBD2-MBD polypeptides immobilized on magnetic beads as described previously [77, 82].

The resulting enriched methylated fraction and the total input fraction were then subjected to library amplification using the NEBNext amplification module according to the manufacturer's protocols. Library fragments that were between 200–300 bp were size selected after agarose gel electrophoresis. The libraries were then subjected to emulsion PCR and bead enrichment following the SOLiD emulsion PCR protocol (Life

Technologies). The resulting beads were then deposited on the SOLiD flow cell and subjected to massively parallel 50 bp single-read sequencing on a SOLiD v4.0 sequencer octet segment, with one octet segment for the total input and another one for the enriched methylated fraction. The resulting reads were mapped to hg18, and peaks of reads were called using the MACS program [83]. This allowed us to identify the methylated regions in each sample.

Figure 2.1: ***MBD-Seq work-flow.*** Libraries of DNA samples from HCT116 cell lines are prepared and split into two fractions: one for methylation specific enrichment and another for total input. Methylated fragments were enriched using MBD2 methyl-binding domain polypeptides immobilized on magnetic beads. The methylated DNA is subsequently eluted and subjected to next generation sequencing along with the total input.



2.2.5 RNA-Sequencing

Total RNA was extracted from HCT116 isogenic cell lines using the RNeasy Kit from Qiagen. Approximately 1μg of total RNA from each cell line was prepared for sequencing using the TruSeq Stranded Total RNA Sample Preparation kit (Illumina) following manufacturer's protocols. The resulting cDNA libraries were run on the Illumina HS200 platform for 200 cycles with 100bp by 100bp paired end reads. Reads were mapped to hg18 using RSEM.

2.2.6 Analysis of Differentially Methylated Regions

BAM files generated from MBD-seq were loaded into the Integrative Genomic Viewer [84, 85] to visualize the location of methylated peaks in the genome. CEAS software [86, 87] was used to calculate the fraction of identified methylated and differentially methylated regions overlapping with various genome annotations (introns, exons, 5' gene upstream, 3' gene downstream, distal intergenic).

2.2.6 Bisulfite Sequencing

Bisulfite sequencing was used to detect the presence of methylated genomic regions of interest. This method involves conversion of unmethylated cytosines on CpG sites to uracils, while methylated cytosines remain unchanged. After PCR amplification, uracils are read as thymidines, and all cytosines present represent methylated cytosines.

Genomic DNA from a set of isogenic cell lines (HCT116, HCT116-*DNMT1-/-*, HCT116-*DNMT3b-/-*, HCT116-*DNMT1-/- DNMT3b-/-*) was bisulfite converted using the EZ DNA Methylation Gold kit (Zymo). Converted DNA was subject to PCR amplification with a set of primers designed to genomic CpG islands of interest. The following cycling conditions were used: 95 °C for 3 minutes, 38 cycles of 95 °C for 30 seconds, 56 °C for 30 seconds, 72 °C for 45 seconds, and 72 °C for 3 minutes. PCR products were run on 2% agarose gel and excised and purified using the Zymoclean Gel DNA Recovery Kit (Zymo). After recovery, the DNA was cloned into TOP 10 competent cells using the TOPO TA cloning kit (Invitrogen) and manufacturer's protocol. Between 50-100µl of cells were plated on Kanamycin-containing plates. Positive clones were selected for Sanger sequencing. The data obtained were used to compare methylation differences among the HCT116 wild type and knock out cell lines.

2.2.7 Chromatin Immunoprecipitation (ChIP)

Localization of DNMT1 to genomic regions of interest was determined by chromatin immunoprecipitation. HCT116-WT, HCT116-*DNMT1-/-*, HCT116-*DNMT3b-/-*, HCT116-DKO cells are treated with 1% formaldehyde to crosslink chromatin proteins to DNA. Cells were lysed and sonicated to shear DNA into small 300bp fragments. A small sample was set aside as the input fraction. Antibodies against DNMT1 (Sigma) was applied to the remaining sonicated sample and allowed to incubate overnight at 4°C. 30µL per sample of Protein G Dynabeads (Invitrogen) was blocked in PBS BSA (5mg/mL), supplemented with yeast tRNA, and incubated overnight at 4°C. The next

day, the blocked bead solution was combined with the antibody-DNA solution and incubated for 4 hours. Bead bound immunocomplexes were precipitated by magnet and washed. Isolated antibody-chromatin complexes and the input fraction (non-immunoprecipiated materials) were treated to remove the crosslink and the DNA is purified. Quantitative PCR with primers against CpG island regions of interest was used to amplify and detect the DNMT1 bound DNA.

2.2.8 Quantitative RT-PCR

RNA was isolated from all 4 HCT116 cell lines with the Qiagen All-Prep kit. 1μg of RNA was converted to cDNA with the High Capacity cDNA Reverse Transcription kit (Applied Biosystems). 10ng/μL of the resulting cDNA was used as the template for real-time PCR. Gene expression was detected by quantitative RT-PCR using EVA Green qPCR master mix (Biotium) on Biorad iCycler thermal cyclers under the following conditions: 1 x 95°C for 10 mm; 45 x 95°C for 30 s, 60°C for 30 s, 72°C for 30 s.

2.3 RESULTS

2.3.1 Genomic DNA methylation patterns established with MBD-sequencing

In order to examine the role of DNMT1 and DNMT3b in establishing/maintaining genome-wide methylation patterns, we used a set of HCT116 isogenic cell lines in which DNMT1, DNMT3b, or both or neither were genetically disrupted [59, 60] (figure 2.2). HCT116, a colorectal carcinoma line, exhibits hypermethylation associated gene silencing at hundreds of genes. DNA was isolated from each of the 4 lines (HCT116, HCT116-*DNMT1-/-*, HCT116-*DNMT3b-/-*, HCT116-*DNMT1-/- DNMT3b-/-*) and sequenced by MBD-seq. CEAS software [86, 87] was used to display methylation profiles and enrichment at important genomic features. The distribution of methylated regions by chromosome was generated for all 4 isogenic HCT116 cell lines. While HCT116-WT, HCT116-*DNMT1-/-* and HCT116-*DNMT3b-/-* cells showed similar levels of methylation across the genome, a substantial drop in methylated peaks was observed in HCT116-DKO (figures 2.3-2.6). Comparing the percent of methylated regions in each chromosome to the genome background (hg18), we observed that methylated regions are enriched in chromosomes 16, 17, 19, and 22 across all HCT116 cell lines. Additionally, methylated regions in HCT116-DKO were also enriched in chromosome 6. The distribution of methylated regions across important genomic features is depicted in a series of pie charts. While 8-9% of methylated regions occur in promoter sites of HCT116-WT, HCT116-*DNMT1-/-* and HCT116-*DNMT3b-/-*, only 4% are present in

promoter sites of HCT116-DKO. These data indicate that DNMT1 and/or DNMT3b are needed to maintain methylation at gene promoters.

In the human genome, non-coding introns and intergenic regions account for >90% of sequences. Approximately 2.5% of the genome occurs in promoter sites, and 1.9% in coding exons. Compared to this genomic background, HCT116-WT cells were significantly enriched for methylation at promoter and exonic regions (Table 2.1), and relatively depleted for methylation in intronic and distal intergenic regions. Similar changes were observed in HCT116-*DNMT1-/-* and HCT116-*DNMT3b-/-* (Table 2.2, 2.3). In HCT116-DKO, methylated regions were mostly enriched in coding exons (Table 2.4).

Figure 2.2. ***Western Blot of HCT116 Isogenic Cell Lines.***

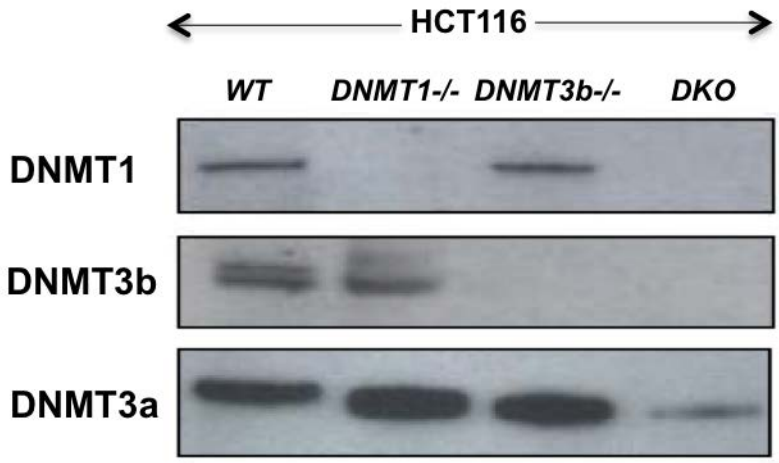


Figure 2.3: ***Methylation in HCT116-WT.***

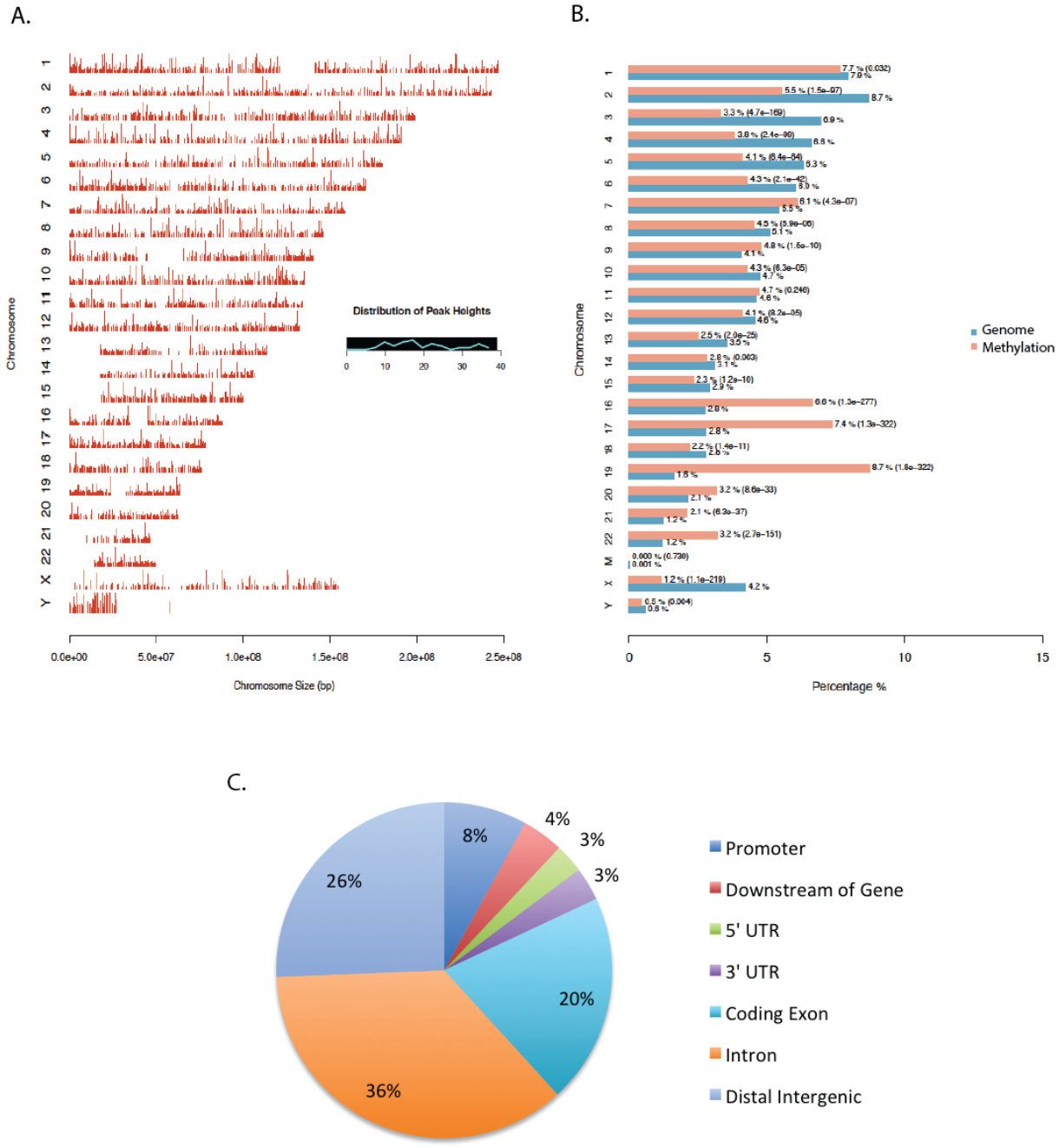


Figure 2.3: ***Methylation in HCT116-WT.*** (A) Distribution of methylated regions by chromosome. The red bars plot methlylation while the line graph illustrates the distribution of peak heights. (B) Percent of methylated regions in each chromosome. The blue bars represent the percentages of the whole mappable regions in the chromosome (genome background). The orange bars represent the percent of methylated regions as a fraction all methylated regions identified in HCT116-WT. P-values for the significance of the relative enrichment of methylated regions with respect to the genome background are shown in parenthesis next to the percentages of the orange bars. The sum of the blue bars or the sum of the orange bars is 100%. (C) Distribution of methylated regions across important genomic features.

Figure 2.4: ***Methylation in HCT116-DNMT1-/-.***

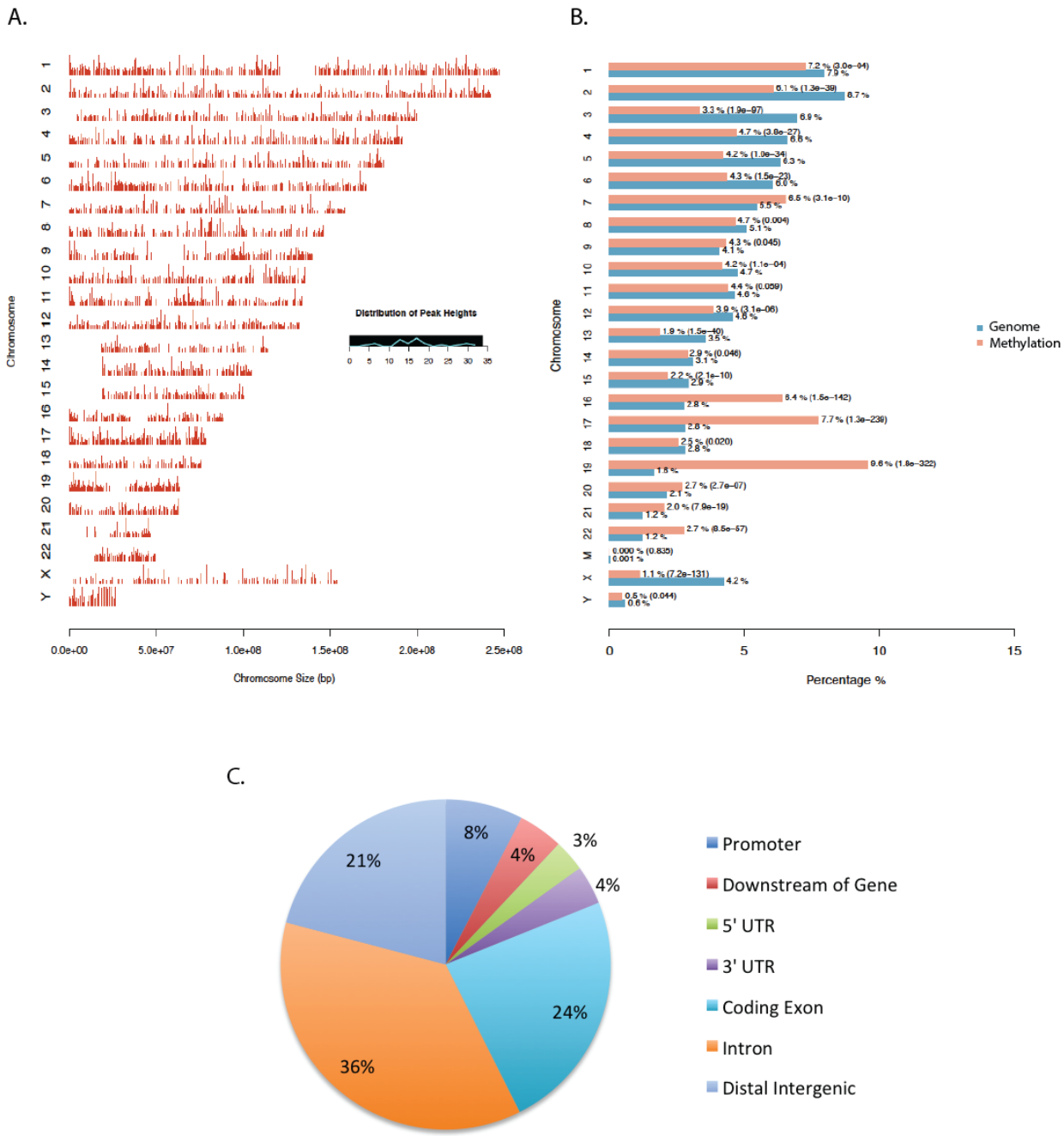


Figure 2.4: ***Methylation in HCT116-DNMT1-/-.*** (A) Distribution of methylated regions by chromosome. The red bars plot methlylation while the line graph illustrates the distribution of peak heights. (B) Percent of methylated regions in each chromosome. The blue bars represent the percentages of the whole mappable regions in the chromosome (genome background). The orange bars represent the percent of methylated regions as a fraction all methylated regions identified in HCT116-*DNMT1-/-*. P-values for the significance of the relative enrichment of methylated regions with respect to the genome background are shown in parenthesis next to the percentages of the orange bars. The sum of the blue bars or the sum of the orange bars is 100%. (C) Distribution of methylated regions across important genomic features.

Figure 2.5: ***Methylation in HCT116-DNMT3b-/-.***

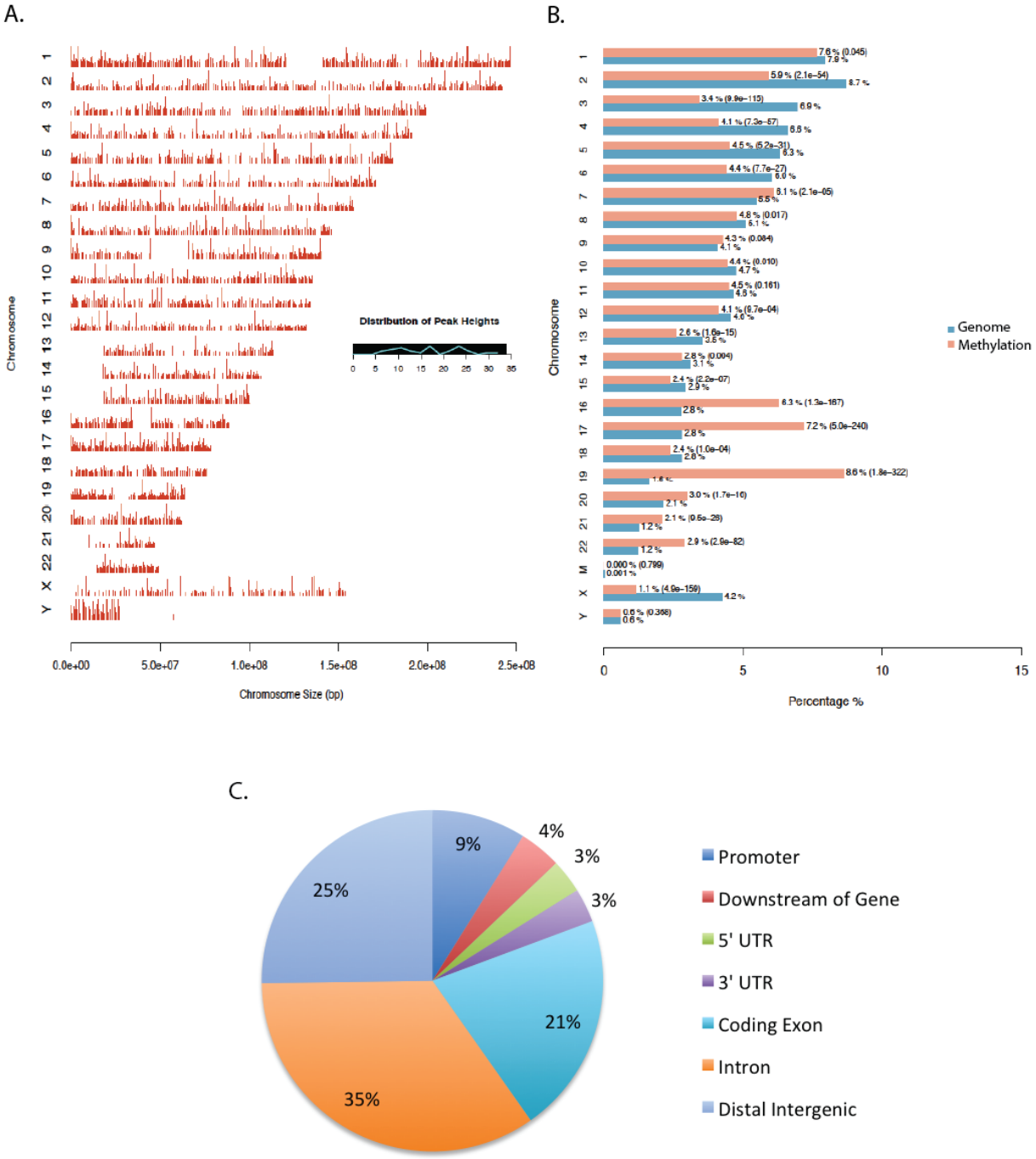


Figure 2.5: ***Methylation in HCT116-DNMT3b-/-.*** (A) Distribution of methylated regions by chromosome. The red bars plot methlylation while the line graph illustrates the distribution of peak heights. (B) Percent of methylated regions in each chromosome. The blue bars represent the percentages of the whole mappable regions in the chromosome (genome background). The orange bars represent the percent of methylated regions as a fraction all methylated regions identified in HCT116-*DNMT3b-/-*. P-values for the significance of the relative enrichment of methylated regions with respect to the genome background are shown in parenthesis next to the percentages of the orange bars. The sum of the blue bars or the sum of the orange bars is 100%. (C) Distribution of methylated regions across important genomic features.

Figure 2.6: ***Methylation in HCT116-DKO.***

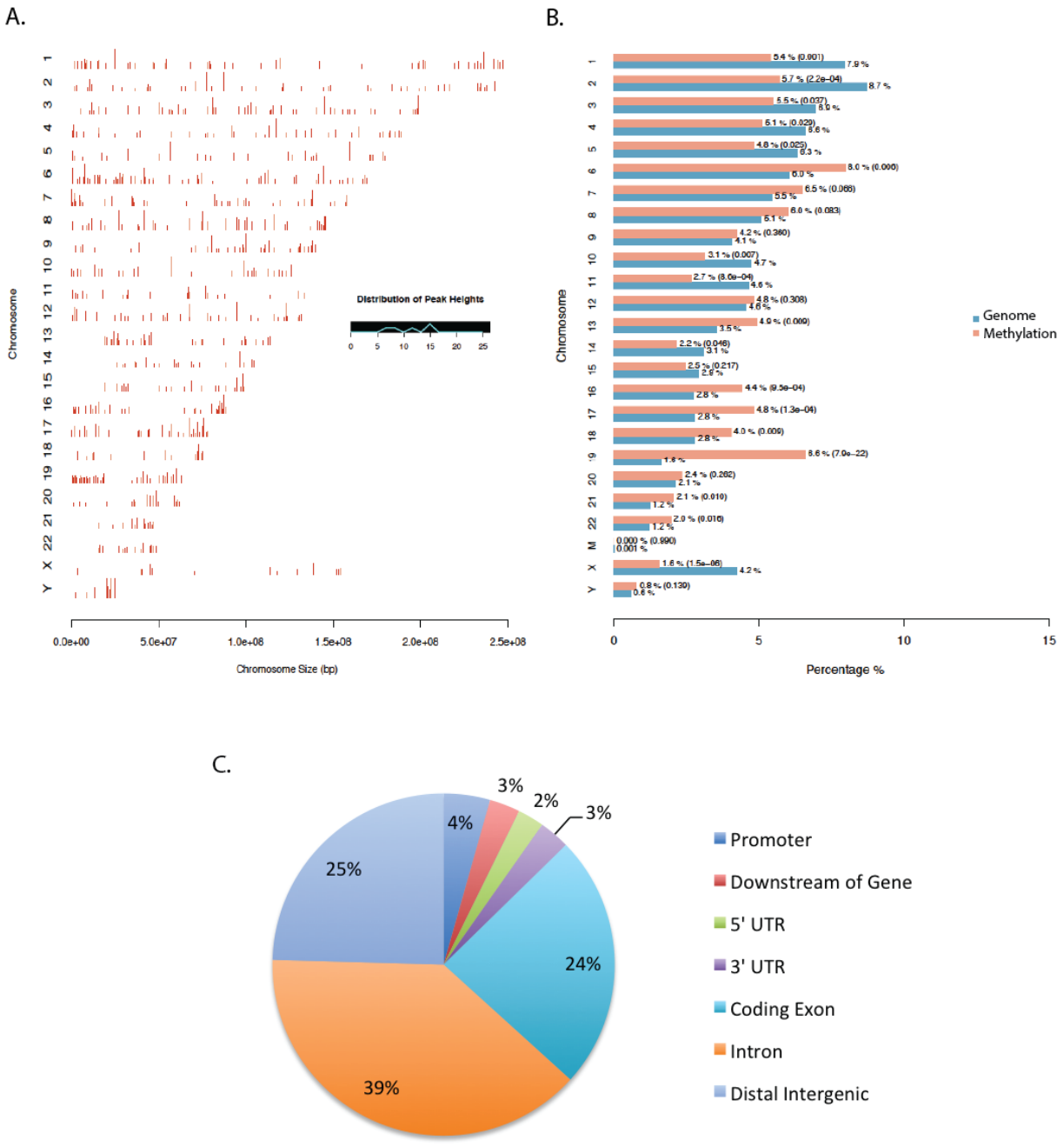


Figure 2.6: ***Methylation in HCT116-DKO.*** (A) Distribution of methylated regions by chromosome. The red bars plot methlylation while the line graph illustrates the distribution of peak heights. (B) Percent of methylated regions in each chromosome. The blue bars represent the percentages of the whole mappable regions in the chromosome (genome background). The orange bars represent the percent of methylated regions as a fraction all methylated regions identified in HCT116-*DKO.* P-values for the significance of the relative enrichment of methylated regions with respect to the genome background are shown in parenthesis next to the percentages of the orange bars. The sum of the blue bars or the sum of the orange bars is 100%. (C) Distribution of methylated regions across important genomic features.

Table 2.1: ***Distribution of Methylated Regions in HCT116-WT Across Important Genomic Features.*** For each genomic feature, the percent of methylated regions in HCT116-WT is shown under the percent of methylated regions expected for the genome (hg18). The percent of methylated regions in each genomic feature were also calculated for the top 500 regions as well as those with p-values < $10^{\wedge -10}$. 'Distal Intergenic' refers to regions that do not belong in any of the other genomic features.

	# of Meth. Regions	Promoter	Downstream Of Gene	5'UTR	3'UTR	Coding Exon	Intron	Distal Intergenic
Genome		2.4%	2.2%	0.4%	1.4%	1.9%	41.6%	50%
HCT116 WT (Total)	29,242	8%	4%	2.8%	3.2%	20.2%	36.1%	25.6%
HCT116 WT (p-value < 10^{-10})	16,033	9.8%	4.3%	3%	4.2%	24.2%	32.5%	22%
HCT116 WT (Top 500)	500	11.6%	5.4%	3.5%	5.1%	30.4%	28.8%	15.2%

Table 2.2: ***Distribution of Methylated Regions in HCT116-DNMT1-/- Across Important Genomic Features.*** For each genomic feature, the percent of methylated regions in HCT116-*DNMT1-/-* is shown under the percent of methylated regions expected for the genome (hg18). The percent of methylated regions in each genomic feature were also calculated for the top 500 regions as well as those with p-values < 10^-10. 'Distal Intergenic' refers to regions that do not belong in any of the other genomic features.

	# of Meth. Regions	Promoter	Downstream Of Gene	5'UTR	3'UTR	Coding Exon	Intron	Distal Intergenic
Genome		2.4%	2.2%	0.4%	1.4%	1.9%	41.6%	50.0%
HCT116-DNMT1-/- (Total)	16,661	7.6%	4.4%	3.1%	3.8%	23.7%	36.5%	20.9%
HCT116-DNMT1-/- (p-value < 10^{-10})	8870	8.6%	4.7%	3.9%	4.1%	26.5%	33.9%	18.3%
HCT116-DNMT1-/- (Top 500)	500	9.7%	4.8%	3.2%	4.9%	33.1%	29.3%	15.0%

Table 2.3: ***Distribution of Methylated Regions in HCT116-DNMT3b-/- Across Important Genomic Features.*** For each genomic feature, the percent of methylated regions in HCT116-*DNMT3b-/-* is shown under the percent of methylated regions expected for the genome (hg18). The percent of methylated regions in each genomic feature were also calculated for the top 500 regions as well as those with p-values < 10^-10. 'Distal Intergenic' refers to regions that do not belong in any of the other genomic features.

	# of Meth. Regions	Promoter	Downstream Of Gene	5'UTR	3'UTR	Coding Exon	Intron	Distal Intergenic
Genome		2.4%	2.2%	0.4%	1.4%	1.9%	41.6%	50.0%
HCT116-DNMT3b-/- (Total)	20,804	8.9%	4.0%	3.2%	3.2%	20.9%	34.5%	25.2%
HCT116-DNMT3b-/- (p-value < 10^{-10})	10,489	10.7%	4.0%	4.0%	3.7%	23.3%	32.2%	22.0%
HCT116-DNMT3b-/- (Top 500)	500	11.0%	4.7%	3.4%	5.0%	29.2%	28.2%	18.5%

Table 2.4: ***Distribution of Methylated Regions in HCT116-DKO Across Important Genomic Features.*** For each genomic feature, the percent of methylated regions in HCT116-DKO is shown under the percent of methylated regions expected for the genome (hg18). The percent of methylated regions in each genomic feature were also calculated for the top 500 regions as well as those with p-values < $10^{\wedge -10}$. 'Distal Intergenic' refers to regions that do not belong in any of the other genomic features.

	# of Meth. Regions	Promoter	Downstream Of Gene	5'UTR	3'UTR	Coding Exon	Intron	Distal Intergenic
Genome		2.4%	2.2%	0.4%	1.4%	1.9%	41.6%	50.0%
HCT116 DKO (Total)	1015	4.4%	2.9%	2.5%	2.9%	24.1%	38.8%	24.6%
HCT116 DKO (p-value < 10^{-10})	266	4.5%	2.7%	2.3%	2.6%	35.3%	37.2%	15.4%
HCT116 DKO (Top 500)	500	3.8%	3.2%	3.0%	3.2%	31.1%	37.6%	18.1%

2.3.2 DNMT1 and DNMT3b have both distinct and cooperative functions in maintaining methylation patterns.

From MBD-sequencing, we identified over 15,000 regions of differential methylation. In 3233 regions, we found DNMT1-dependent methylation, defined as regions where loss of DNMT1 alone is sufficient to produce loss of methylation marks. These DNMT1-dependent methylated sites are concentrated in putative promoter regions (+/- 3000bp away from TSS) and gene body exons. 10% of the 3233 regions were observed in gene promoters, a value that increases to almost 14% in the top 500 observed hits (table 2.5). DNMT1 dependent regions were also enriched in coding exons, with a 5-fold increase from genome background. Average methylation profiles reveal a substantial drop in methylation around the transcriptional start sites in HCT116-*DNMT1-/-* cells as compared to the HCT116-WT control (figure 2.10). Taken together, these findings suggest that DNMT1 is required for maintaining methylation at a significant number of gene promoters and exons. A list of the top putative promoter regions exhibiting DNMT1 mediated methylation is provided in Table 2.8.

DNMT1 mediates loss of methylation in more regions than DNMT3b, as our results revealed only 1243 regions that displayed DNMT3b dependent methylation. These results can be visualized with methylome maps- revealing more regions of DNMT1 mediated methylation than that of DNMT3b mediated methylation (figures 2.7a, 2.8a). Additionally, the majority of DNMT3b dependent regions are located in coding exons, with 16.5% enrichment in the top 500 hits as compared to the genome background of

1.9% (Table 2.6). DNMT3b does not seem to affect promoter methylation, as the average profile reveals DNMT3b-/- methylation patterns to be similar to that of HCT116-WT (figure 2.11).

Finally, the bulk of differentially methylated regions (13,377) displayed hypomethylation in DKO only, indicating that loss of DNMT1 or DNMT3b alone does not affect the region's hypermethylated state, but absence of both results in loss of methylation (figure 2.9). These regions are enriched in both promoters and coding exons (Table 2.7).

Figure 2.7: ***Loss of Methylation in HCT116-DNMT1-/-.***

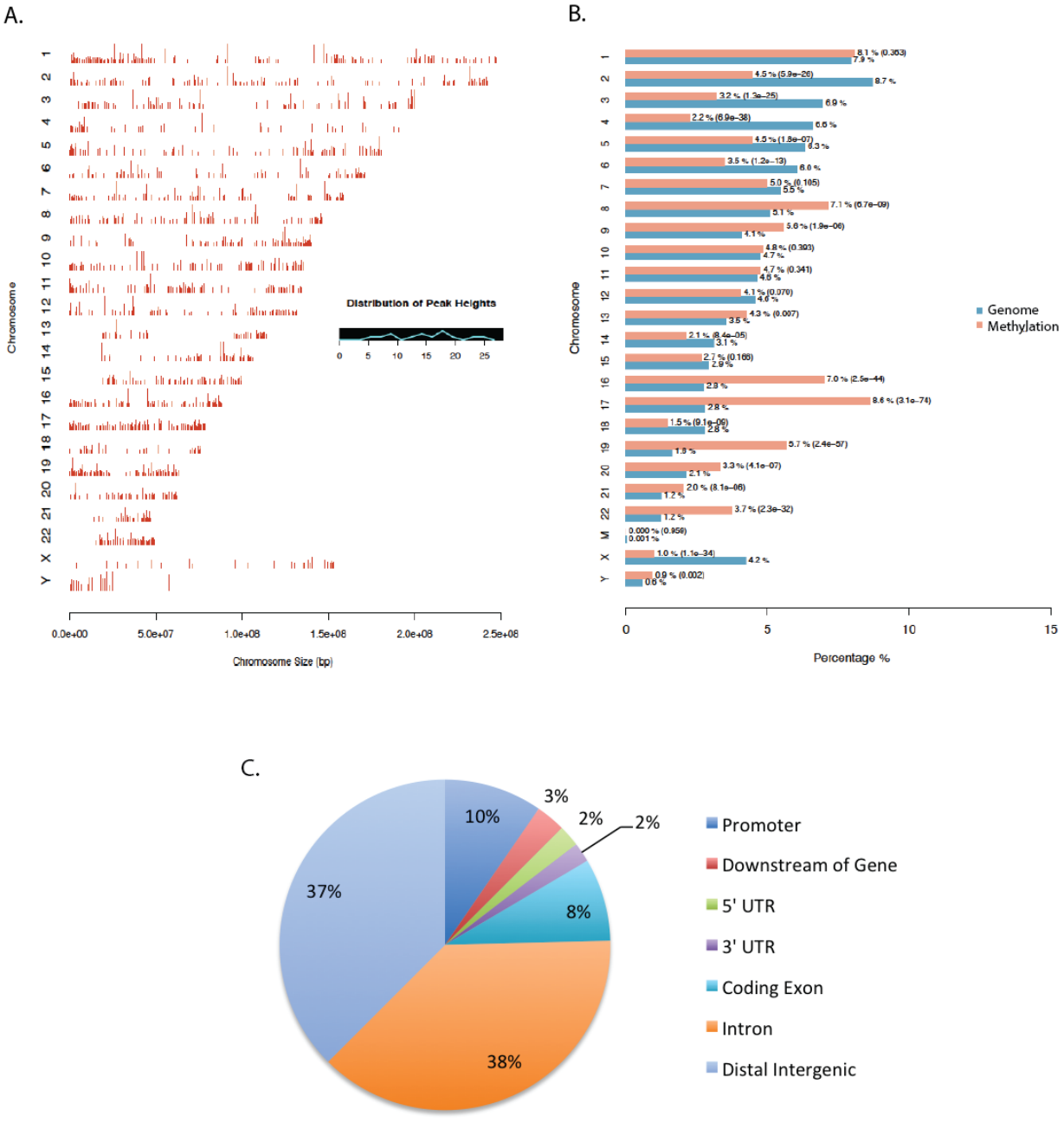


Figure 2.7: ***Loss of Methylation in HCT116-DNMT1-/-.*** Regions that are hypomethylated in HCT116-*DNMT1-/-* relative to the HCT116-WT are displayed. (A) Distribution of differentially methylated regions by chromosome. The red bars plot differentially methlylated regions while the line graph illustrates the distribution of peak heights. (B) Percent of differentially methylated regions in each chromosome. The blue bars represent the percentages of the whole mappable regions in the chromosome (genome background). The orange bars represent the percent of differentially methylated regions as a fraction all methylated regions identified in HCT116-WT. P-values for the significance of the relative enrichment of methylated regions with respect to the genome background are shown in parenthesis next to the percentages of the orange bars. The sum of the blue bars or the sum of the orange bars is 100%. (C) Distribution of differentially methylated regions across important genomic features.

Figure 2.8: ***Loss of Methylation in HCT116-DNMT3b-/-.***

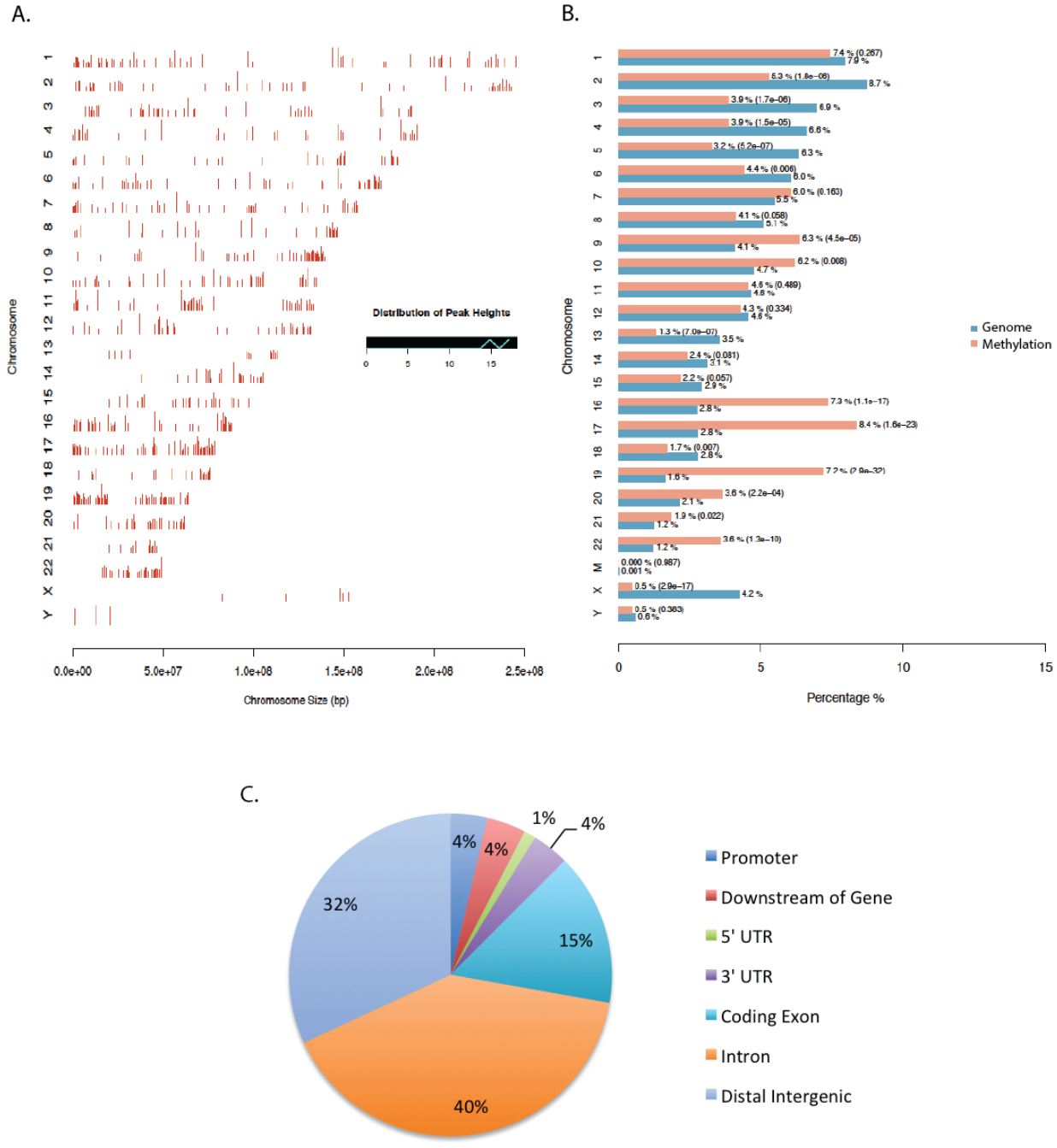


Figure 2.8: ***Loss of Methylation in HCT116-DNMT3b-/-.*** Regions that are hypomethylated in HCT116-*DNMT3b-/-* relative to the HCT116-WT are displayed. (A) Distribution of differentially methylated regions by chromosome. The red bars plot differentially methlylated regions while the line graph illustrates the distribution of peak heights. (B) Percent of differentially methylated regions in each chromosome. The blue bars represent the percentages of the whole mappable regions in the chromosome (genome background). The orange bars represent the percent of differentially methylated regions as a fraction all methylated regions identified in HCT116-WT. P-values for the significance of the relative enrichment of methylated regions with respect to the genome background are shown in parenthesis next to the percentages of the orange bars. The sum of the blue bars or the sum of the orange bars is 100%. (C) Distribution of differentially methylated regions across important genomic features.

Figure 2.9: ***Hypomethylation in HCT116-DKO.***

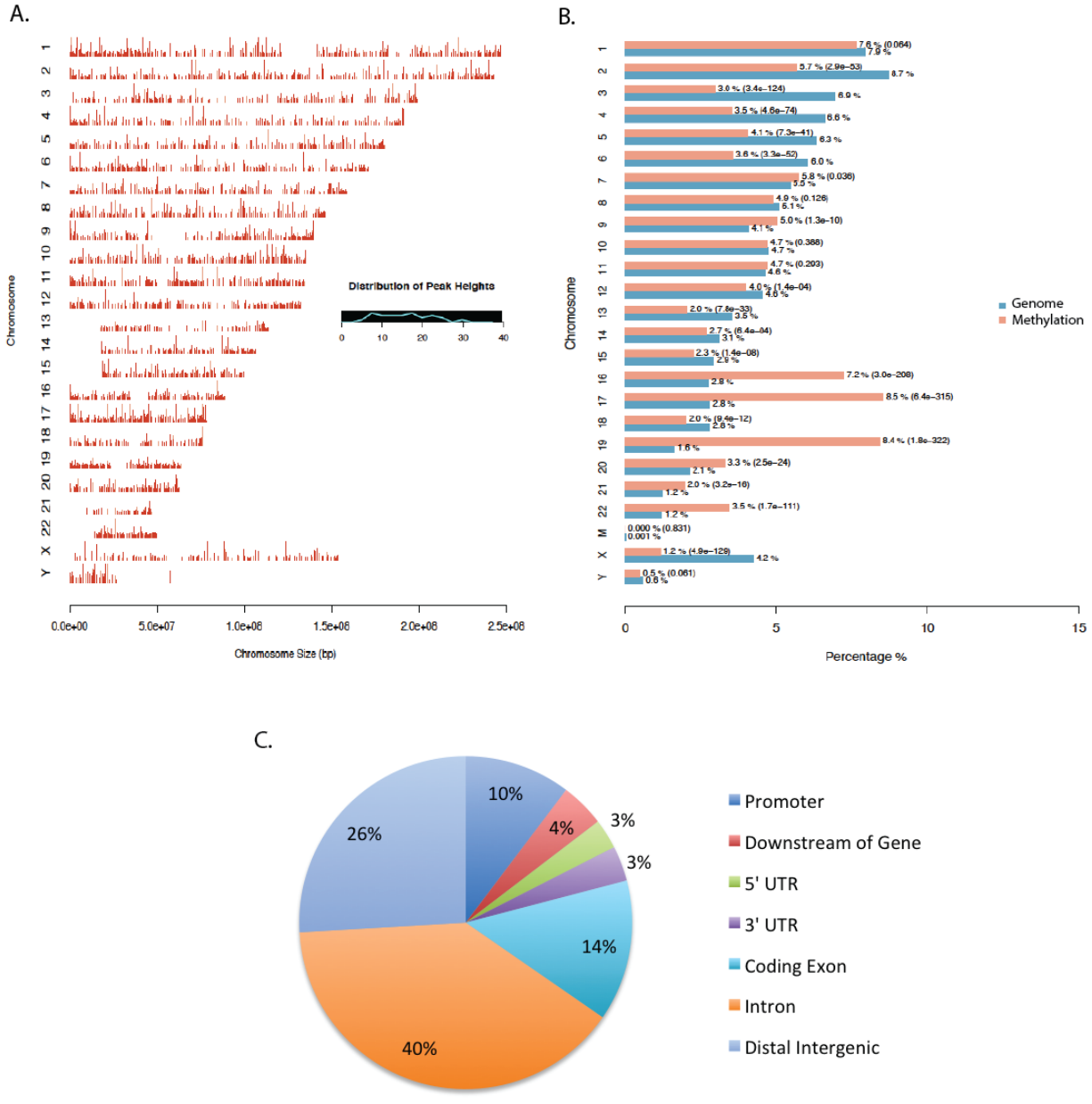


Figure 2.9: ***Hypomethylation in HCT116-DKO.*** Regions hypomethylated in HCT116-DKO but hypermethylated in HCT116-WT, HCT116-*DNMT1-/-* and HCT116-*DNMT3b-/-* are represented. (A) Distribution of hypomethylation by chromosome. The red bars plot hypomethylated regions while the line graph illustrates the distribution of peak heights. (B) Percent of hypomethylated regions in each chromosome. The blue bars represent the percentages of the whole mappable regions in the chromosome (genome background). The orange bars represent the percent of hypomethylated regions as a fraction all methylated regions identified in HCT116-WT. P-values for the significance of the relative enrichment of methylated regions with respect to the genome background are shown in parenthesis next to the percentages of the orange bars. The sum of the blue bars or the sum of the orange bars is 100%. (C) Distribution of hypomethylated regions across important genomic features.

Figure 2.10: ***Average Methylation Profiles (Continued on next page)***

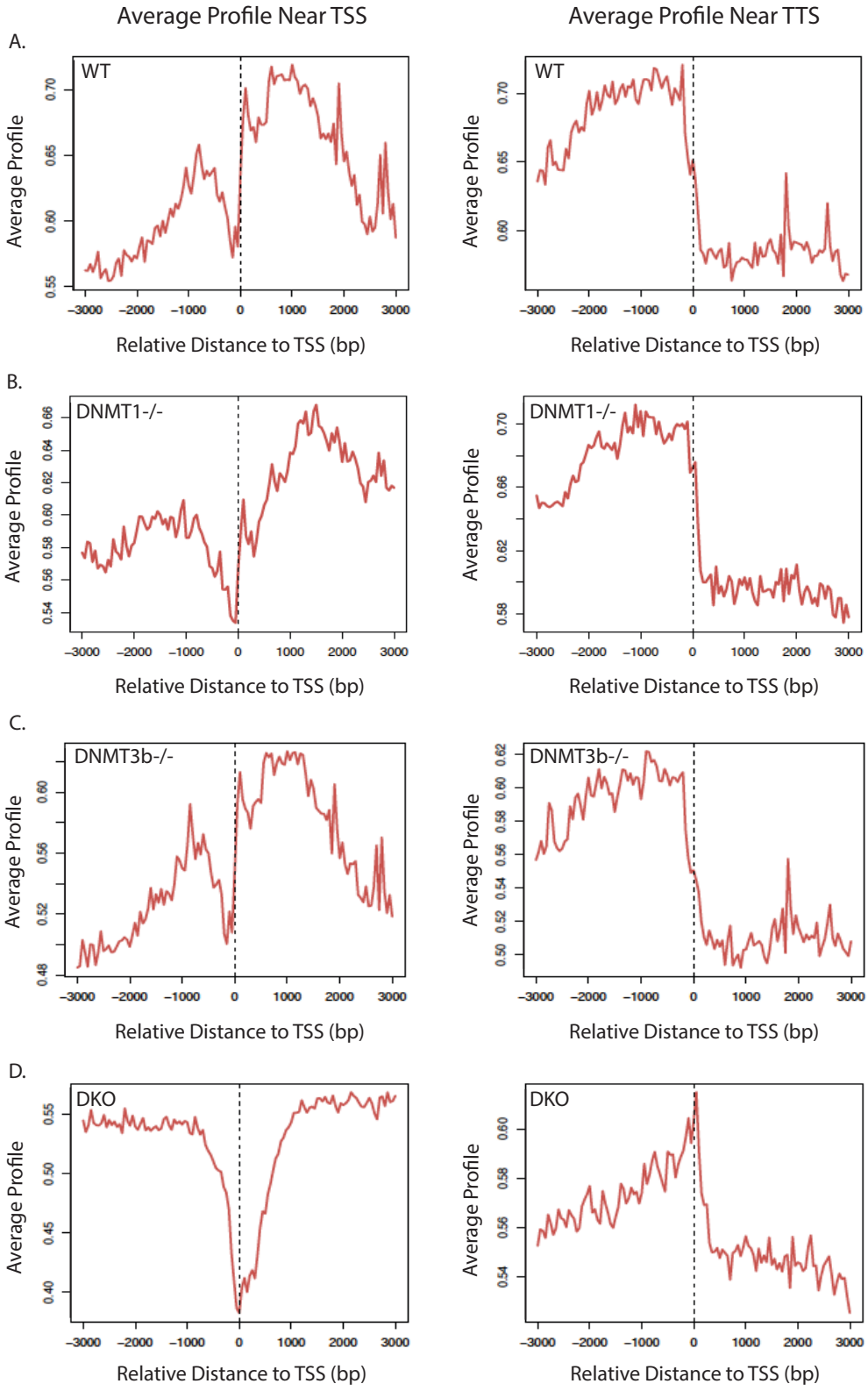


Figure 2.10: ***Average Methylation Profile Near the Transcriptional Start Site (TSS) and Transcriptional Termination Site (TTS) for (A) HCT116-WT, (B) HCT116-DNMT1-/-, (C) HCT116-DNMT3b-/-, and (D) HCT116-DKO.*** Average enrichment of methylated regions is shown for regions near the TSS and TTS. HCT116-DNMT1-/- cells show a slight dip in methylation after the TSS (B) as compared to WT (A). HCT116-DNMT3b-/- cells show similar methylation levels near the TSS (C). HCT116-DKO remains hypomethylated around the TSS (D).

Table 2.5: ***Genomic Distribution of Differential Methylation in Regions Presenting DNMT1 dependent methylation.*** Regions hypermethylated in HCT116-WT but hypomethylated in HCT116-*DNMT1-/-* are represented. For each genomic feature, the percent of differentially methylated regions is shown under the percent of methylated regions expected for the genome (hg18). The percent of differentially methylated regions in each genomic feature were also calculated for the top 500 regions as well as those with p-values < $10^{\wedge -10}$. We observe that regions presenting DNMT1 dependent methylation are enriched in promoter regions and gene body exons (outlined in red).

	Promoter <=3000bp	Downstream Of Gene	5'UTR	3'UTR	Coding Exon	Intron	Distal Intergenic
Genome	2.4%	2.2%	0.4%	1.4%	1.9%	41.6%	50%
Total Observed (3233)	**9.6%**	**2.9%**	**2.1%**	**1.9%**	**8.1%**	**37.9%**	**37.5%**
p-value < 10^{-10} (1177)	**12.1%**	**3.1%**	**2.7%**	**2.1%**	**9.9%**	**33.5%**	**36.5%**
Top Observed (500)	**13.6%**	**3.7%**	**2.3%**	**2.2%**	**10.5%**	**31.1%**	**36.6%**

Table 2.6: ***Genomic Distribution of Differential Methylation in Regions Presenting DNMT3b dependent methylation.*** Regions hypermethylated in HCT116-WT but hypomethylated in HCT116-*DNMT3b-/-* are represented. For each genomic feature, the percent of differentially methylated regions is shown under the percent of methylated regions expected for the genome (hg18). The percent of differentially methylated regions in each genomic feature were also calculated for the top 500 regions as well as those with p-values < $10^{\wedge -10}$. 'Distal Intergenic' refers to regions that do not belong in any of the other genomic features. We observe that regions presenting DNMT3b dependent methylation are enriched in gene body exons (outlined in red).

	Promoter	Downstream Of Gene	5'UTR	3'UTR	Coding Exon	Intron	Distal Intergenic
Genome	2.4%	2.2%	0.4%	1.4%	1.9%	41.6%	50%
Total Observed (1243)	**3.7%**	**3.9%**	**1.2%**	**3.7%**	**15.2%**	**40.3%**	**31.9%**
p-value < 10^{-10} (249)	**3.2%**	**4.9%**	**2.1%**	**4.2%**	**18%**	**35%**	**32.5%**
Top Observed (500)	**3.4%**	**4.5%**	**1.1%**	**4.3%**	**16.5%**	**37.6%**	**32.6%**

Table 2.7: ***Genomic Distribution of Differential Methylation in Regions Hypomethylated in HCT116-DKO.*** Regions hypomethylated in HCT116-DKO but hypermethylated in HCT116-WT, HCT116-*DNMT1-/-* and HCT116-*DNMT3b-/-* are represented. For each genomic feature, the percent of hypomethylated regions is shown under the percent of methylated regions expected for the genome (hg18). The percent of hypomethylated regions in each genomic feature were also calculated for the top 500 regions as well as those with p-values < $10^{\wedge -10}$. 'Distal Intergenic' refers to regions that do not belong in any of the other genomic features.

	Promoter	Downstream Of Gene	5'UTR	3'UTR	Coding Exon	Intron	Distal Intergenic
Genome	2.4%	2.2%	0.4%	1.4%	1.9%	41.6%	50%
Total Observed (13,377)	**10.3%**	**4.3%**	**2.9%**	**3.4%**	**13.7%**	**39.4%**	**25.9%**
p-value < 10^{-10} (8071)	**11.8%**	**4%**	**2.5%**	**4.2%**	**14.1%**	**39%**	**26.4%**
Top Observed (500)	**14.3%**	**4.1%**	**3%**	**4.3%**	**14.9%**	**38.6%**	**26.3%**

Table 2.8: ***Top Putative Promoter Regions Exhibiting DNMT1 Mediated Methylation.***

Gene	Location	TSS Distance	TSS Orientation	P-Value
ALKBH5	chr17:18027236-18030236	1144	downstream	1950.99
HOXB8	chr17:44045954-44048954	153	contains	1531.43
ARC	chr8:143688698-143693800	1587	downstream	1213.7
CBX8	chr17:75386026-75389026	2040	upstream	855.84
EPM2AIP1	chr3:37007938-37010937	363	contains	829.96
PURB	chr7:44889489-44892489	497	contains	812.68
LHX6	chr9:124029007-124032007	334	contains	788.06
NT5DC2	chr3:52542863-52545862	228	contains	611.25
LFNG	chr7:2523317-2526316	793	downstream	557.51
DNAJC22	chr12:48023494-48028650	1236	contains	555.03
SKAP2	chr7:26869711-26872711	344	upstream	539.07
PIM1	chr6:37245643-37248643	1243	downstream	498.34
NTF3	chr12:5472734-5475734	675	downstream	494.73
SKI	chr1:2149391-2152391	897	downstream	493.59
MN1	chr22:26524778-26527777	1210	downstream	493.28
TEAD3	chr6:35572675-35575674	1334	upstream	489.78
SACM1L	chr3:45703535-45706534	724	upstream	483.07
RET	chr10:42891818-42894817	794	downstream	477.16
TACC1	chr8:38763160-38766159	780	downstream	462.51
ROPN1L	chr5:10493882-10496881	372	contains	443.54
SPOCK1	chr5:136860630-136863630	788	downstream	440.06
NPAS4	chr11:65943948-65946947	396	contains	438.72
PRKCE	chr2:45731899-45734899	852	downstream	397.82
GBGT1	chr9:135027574-135030574	49	contains	388.66
TEKT4	chr2:94899959-94902959	500	contains	387.6
ZNF540	chr19:42732715-42735714	66	contains	380.23
MAN1C1	chr1:25815653-25818652	606	downstream	370.42
VIM	chr10:17310010-17313010	1246	downstream	369.82
KIF2B	chr17:49254416-49257416	678	contains	355.29
MEIS2	chr15:35177311-35180310	80	contains	352.03
ICAM1	chr19:10240718-10243717	300	contains	350.12
CRISPLD1	chr8:76057363-76060363	535	contains	344.59
BATF3	chr1:210939171-210942171	720	contains	342.39
CENPB	chr20:3712953-3715952	886	downstream	341.43
ROBO3	chr11:124240032-124243032	1017	contains	338.69
NR3C2	chr4:149583075-149586075	1452	upstream	337.5
SCRT1	chr8:145528743-145531742	510	contains	332.93
SMPD3	chr16:67037985-67040984	427	contains	322.5
CYP26B1	chr2:72229071-72232071	2099	upstream	318.44
HUNK	chr21:32168282-32171282	2283	downstream	280.72
IFFO1	chr12:6533303-6536302	689	contains	269.26

2.3.3 DNMT1 dependent methylation near gene TSS occurs independent of selection.

Reads generated from MBD-sequencing were loaded into the Integrated Genomics Viewer (IGV), a visualization tool for exploration of genomic data sets. Differentially methylated regions across all 4 HCT116 cell lines are represented by peaks of aligned reads in one or more cell lines with absence of peaks in one or more cell lines. Visual inspection of the top regions (p-value $<10^{-10}$) yielded a subset of 500 genes that contained the most convincing DNMT1-dependent methylated regions. These regions were characterized by hypermethylation in DNMT1 containing cell lines and corresponding hypomethylation in DNMT1-/- lines. To substantiate the differential methylation patterns, bisulfite sequencing was performed on the top regions. Eight clones were sequenced for each region, and the results confirm that DNMT1 is necessary for maintaining methylation at these regions (figure 2.11).

Hypermethylation in all cell lines except HCT116-DKO characterized the majority of identified regions. This indicates that DNMT1 and DNMT3b may have similar mechanisms of action in these regions and are likely to act redundantly to maintain methylation. These results were also validated in a subset of regions with bisulfite sequencing (figure 2.12).

A major concern with the use of isogenic cell lines is clonal selection. Since each of these lines was generated from a single clone, the possibility remains that unselected clones

may have a different phenotype. To confirm that our observations were independent of selection, we knocked down DNMT1 in HCT116 via DNMT1 siRNA over a period of 4 days. Isolated DNA from transfected cells was subjected to bisulfite sequencing analysis with primers specific to select regions of DNMT1-dependent methylation and select regions of DKO hypomethylation. The majority of clones in DNMT1 dependent regions showed at least 60% reduction in CpG methylation upon depletion of DNMT1, while clones in regions of DKO hypomethylation mostly retained their methylated status (figure 2.13). Taken together, our results indicate that DNMT1 dependent methylation in HCT116 occurs independent of selection.

Figure 2.11: ***Representative Regions Exhibiting DNMT1 Mediated Methylation.*** MBD-seq revealed regions of differential methylation regulated by DNMT1. Data for two representative genes, EPM2AIP1 and CYP26B1, are shown here. The blue peaks represent hypermethylated regions in HCT116-WT and HCT116-*DNMT3b-/-*. Bisulfite sequencing of a sub-region displaying DNMT1 dependent methylation (red bar) revealed methylation of almost all CpGs in the presence of DNMT1 and corresponding hypomethylation in the absence of DNMT1.

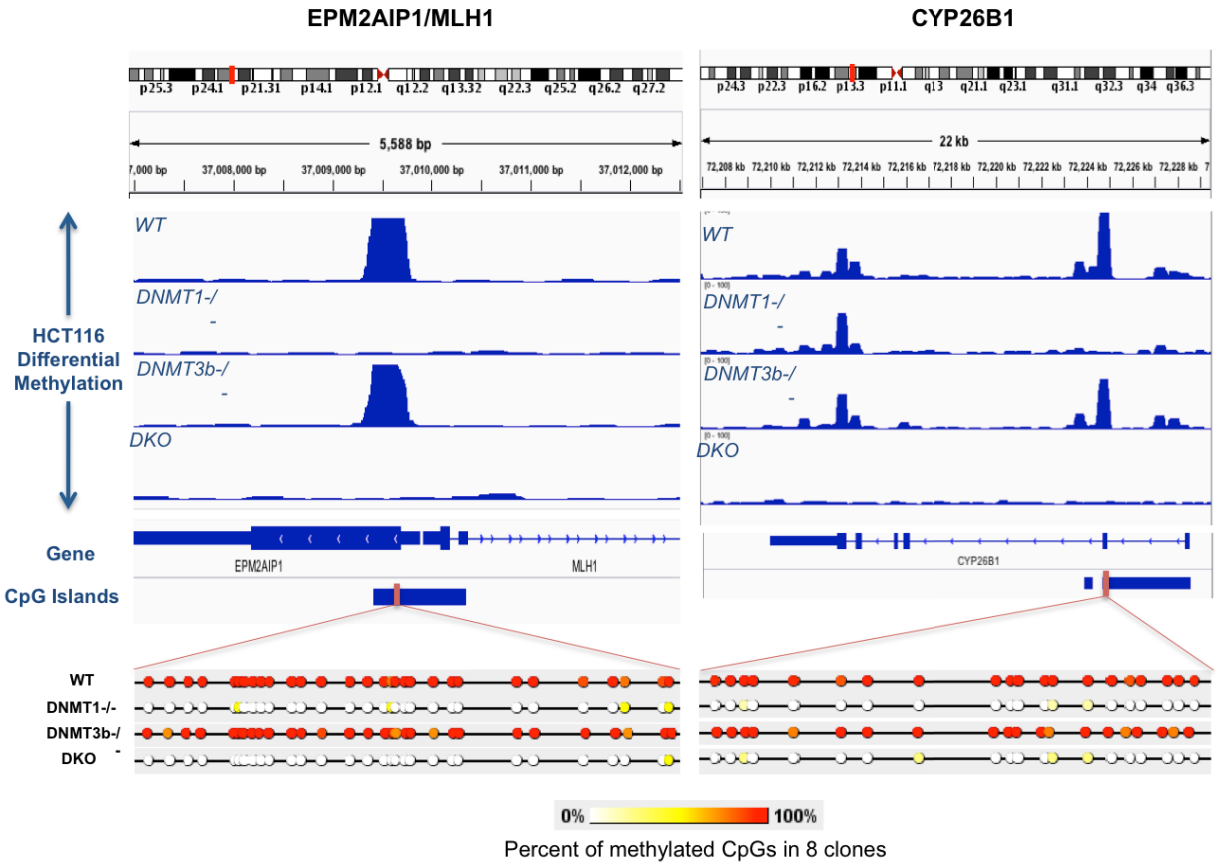


Figure 2.12: ***Representative Regions Exhibiting Hypomethylation in HCT116-DKO.*** MBD-seq results revealed numerous regions with hypomethylated CpG islands in HCT116-DKO only. Data for two representative genes, TBX3 and FLNC, are shown here. Blue peaks indicate the presence of hypermethylated regions in HCT116-WT, HCT116-*DNMT1-/-*, and HCT116-*DNMT3b-/-*. Bisulfite sequencing of 8 representative clones per region confirmed these observations.

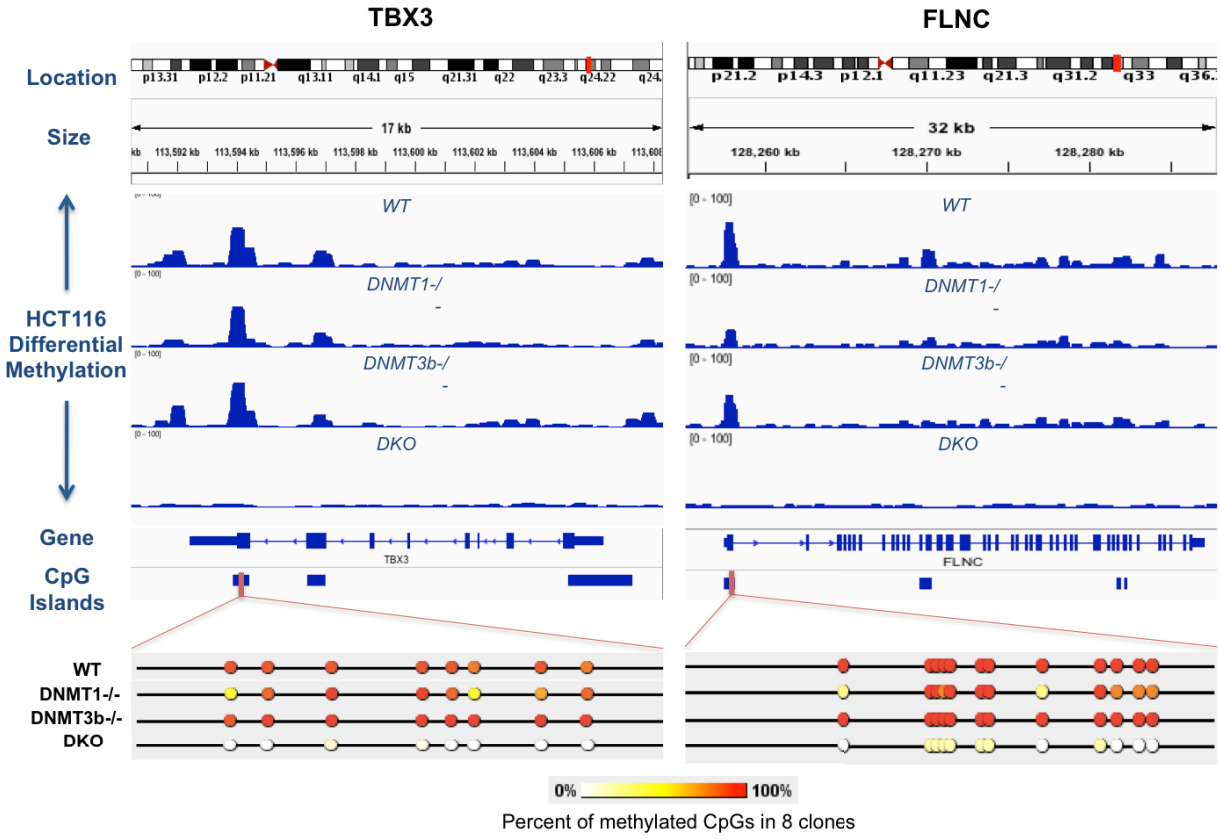
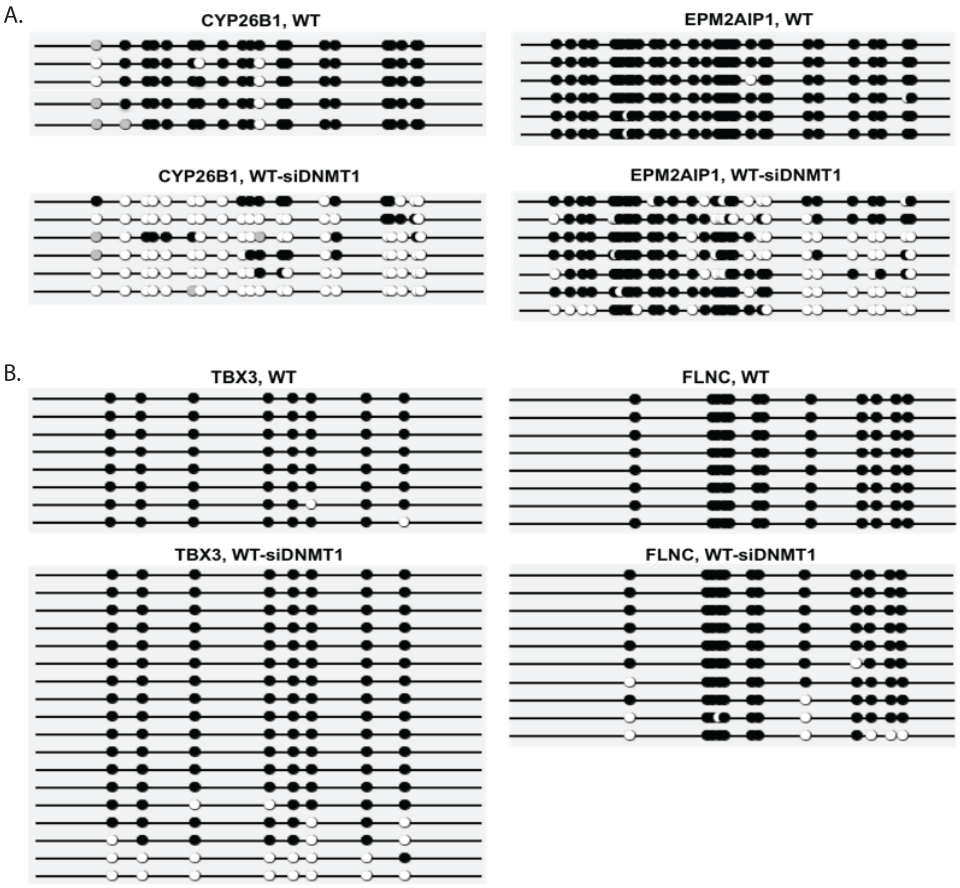


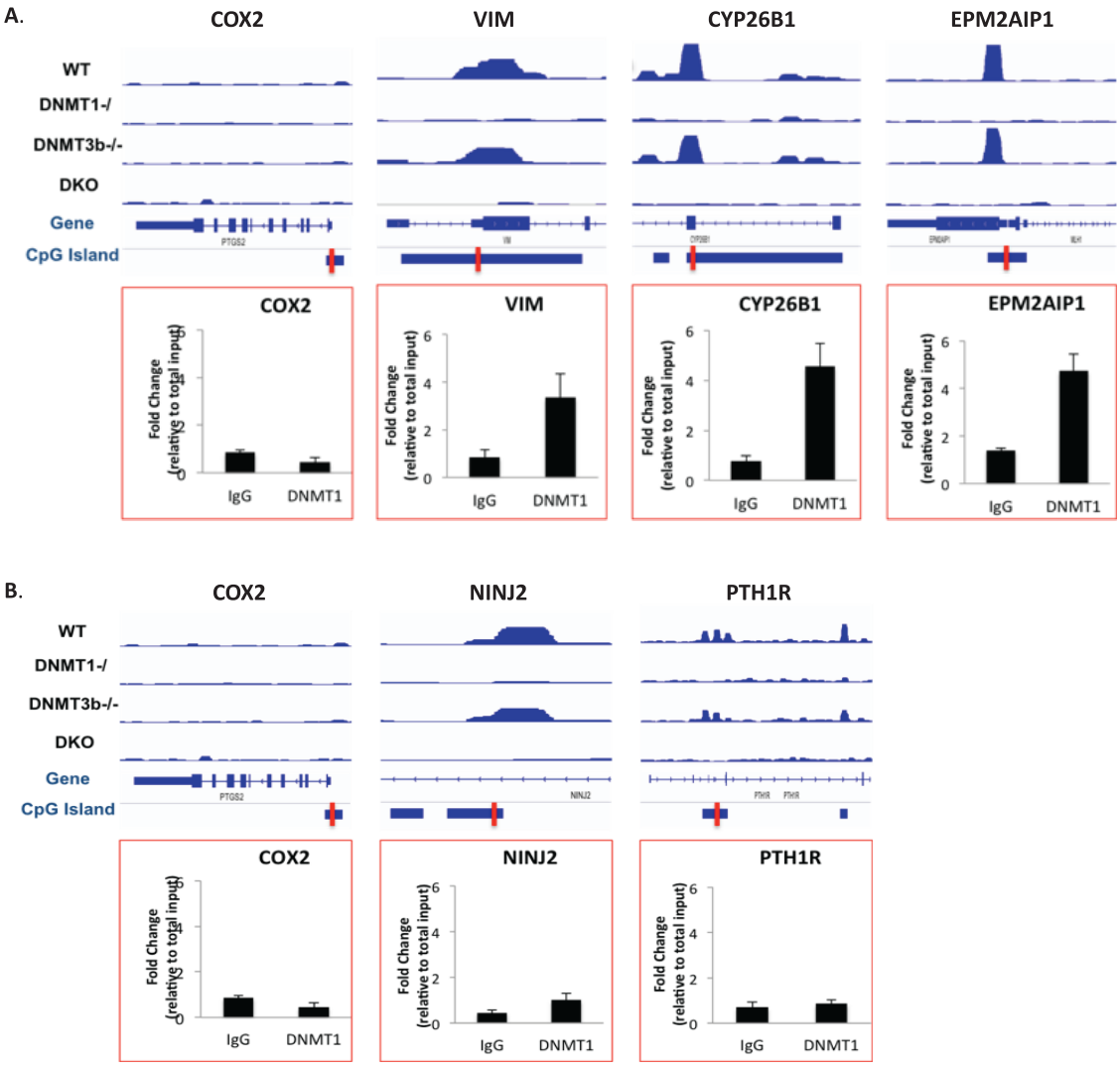
Figure 2.13: ***DNMT1 Mediated Methylation Occurs Independent of Selection.*** (A) Bisulfite sequencing results for representative regions exhibiting DNMT mediated methylation are depicted. Sequenced clones from HCT116-WT and HCT116-WT + siDNMT1 are shown. The black circles denote methylated CpGs while the white circles represent unmethylated CpG. For representative genes CYP26B1 and EPM2AIP1, treatment with siDNMT1 results in a significant reduction in methylated CpG sites. (B) Bisulfite sequencing results for regions exhibiting hypomethylation in HCT116-DKO. The methylation status of representative genes TBX3 and FLNC is mostly unaffected by siDNMT1.



2.3.4 DNMT1 is directly bound to a subset of DNMT1 requiring regions

Next we investigated whether DNMT1 may mediate methylation by binding to targeted regions. Chromatin immunoprecipitation for DNMT1 was performed on regions exhibiting DNMT1 mediated methylation. Results reveal that DNMT1 is bound to CpG islands in putative promoter regions in gene body exons but not bound to CpG islands located at introns (figure 2.14). These results suggest that DNMT1 acts directly on important genomic regions to maintain local methylation marks.

Figure 2.14: ***DNMT1 is Directly Bound to Putative Promoter Regions.*** Select regions exhibiting DNMT1 mediated methylation are represented. (A) Chromatin immunoprecipitation shows DNMT1 directly bound to putative promoter regions. (B) DNMT1 occupancy is not enriched at NINJ2 and PTH1R intronic CpG islands that showed loss of methylation in a DNMT1 dependent manner. An unmethylated region in the COX2 promoter was used as a negative control.



2.3.5 DNMT1 dependent methylation is associated with gene repression

Finally, we investigated whether DNMT1 mediates gene repression directly through DNA methylation. Gene expression profiles for all 4 HCT116 cell lines generated through RNA-seq were compared to the differentially methylated regions identified by MBD-seq. Table 2.9 displays a chart of the results. Among 77 differentially expressed genes identified through RNA-seq, 14 genes showed a negative correlation between methylation and expression. Genes were expressed in the absence of DNMT1 mediated methylation and repressed when DNMT1 was present, suggesting that DNMT1-dependent methylation is associated with gene silencing. This is further supported by the fact that 78% of these regions are located at putative promoter sites, defined as any region within 3000bp of the transcriptional start site.

Further analysis by quantitative RT-PCR confirmed our initial observations. Vimentin, one of our top hits, showed a 10-fold increase in expression in the absence of DNMT1. This result was corroborated by qRT-PCR, with HCT116-*DNMT1-/-* cells displaying an almost 60 fold change in expression when compared to housekeeping gene TBP. When DNMT1 in HCT116 was depleted through siRNA over a period of 7 days, Vimentin expression slowly recovered. Similarly, treatment of HCT116 cells with Decitabine showed a positive correlation between Decitabine concentration used and Vimentin gene expression (figure 2.15). All 14 genes identified displayed similar patterns of methylation dependent expression (figure 2.16). Therefore, these results as a whole provide strong

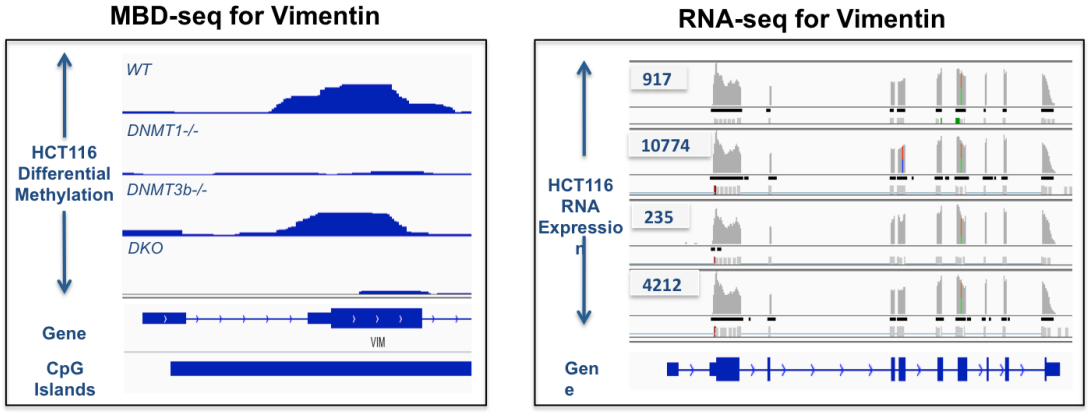
evidence for our hypothesis that DNMT1 requiring methylation can mediate epigenetic gene repression.

Table 2.9: ***DNMT1 Mediated Methylation Can Control Gene Expression.*** Among 77 differentially expressed genes identified through RNA-Seq, 14 regions showed a negative correlation with differential methylation. Differential methylation upstream (+) or downstream (-) of the transcriptional start site (TSS) is notated along with corresponding p-values and number of RNA-Seq reads. Additionally, approximately 78% (11 out of 14 regions) of the regions are present in putative promoter sites (defined as residing within 3000bp of the TSS).

Gene	Differential Methylation						Differential Expression							
	(+) upstream, (-) downstream				Distance	P-value	(✔ indicates expression)				(Number of Reads)			
	wt	1-/-	3b-/-	dko	To TSS		wt	1-/-	3b-/-	dko	wt	1-/-	3b-/-	dko
EPM2AIP1	+		+		363	829.96		✔		✔	34	112	21	83
VIM	-		-		1246	369.82		✔		✔	917	10774	235	4214
ARC	-		-		1587	1213.7		✔		✔	74	109	43	151
CYP26B1	-		-		2099	318.44		✔		✔	40	101	34	70
PTH1R	-		-		4839	314.85		✔		✔	10	37	10	33
NT5DC2	+		+		228	611.25		✔		✔	203	303	195	358
ICAM1	-		-		300	350.12		✔		✔	111	234	10	368
BATF3	+		+		720	342.39		✔		✔	60	133	41	95
IFFO1	-		-		689	269.26		✔		✔	10	29	10	41
NINJ2	-		-		19882	339.56		✔		✔	10	41	10	36
SAP25	+		+		3327	178.97		✔		✔	16	527	10	1303
GBGT1	+		+		49	388.66		✔		✔	10	10	10	34
HUNK	-		-		2283	280.72		✔		✔	10	10	10	36
PURB	+		+		497	812.68		✔		✔	113	184	110	142

Figure 2.15: ***DNMT1 Requiring Methylation Mediates Epigenetic Repression in Vimentin.*** (A) Loss of DNMT1 is associated with corresponding Vimentin expression as shown in MBD-seq and RNA-seq results. (B) Quantitative RT-PCR confirms DNMT1 mediated gene repression. Transient silencing of DNMT1 through RNAi and Decitabine recovers Vimentin expression.

A.



B.

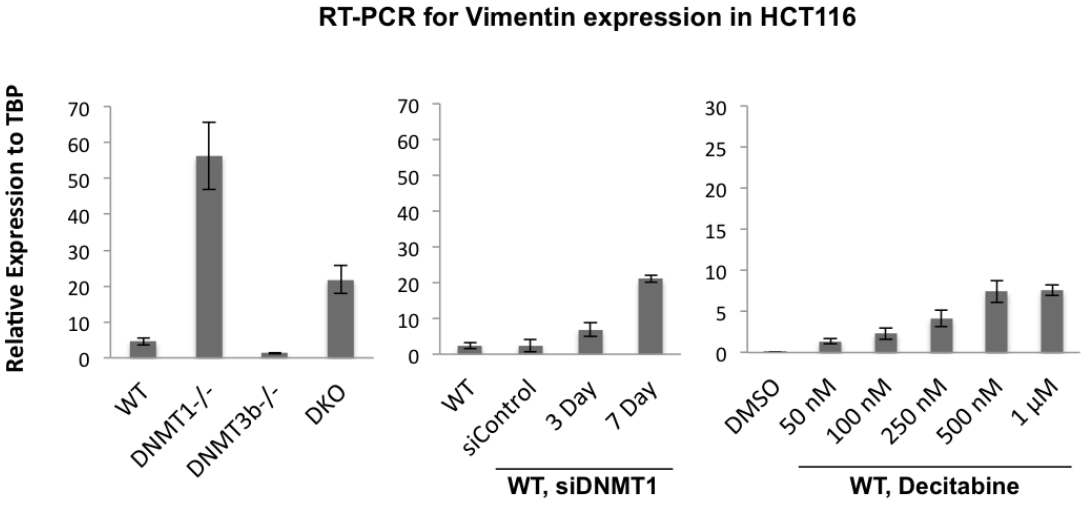
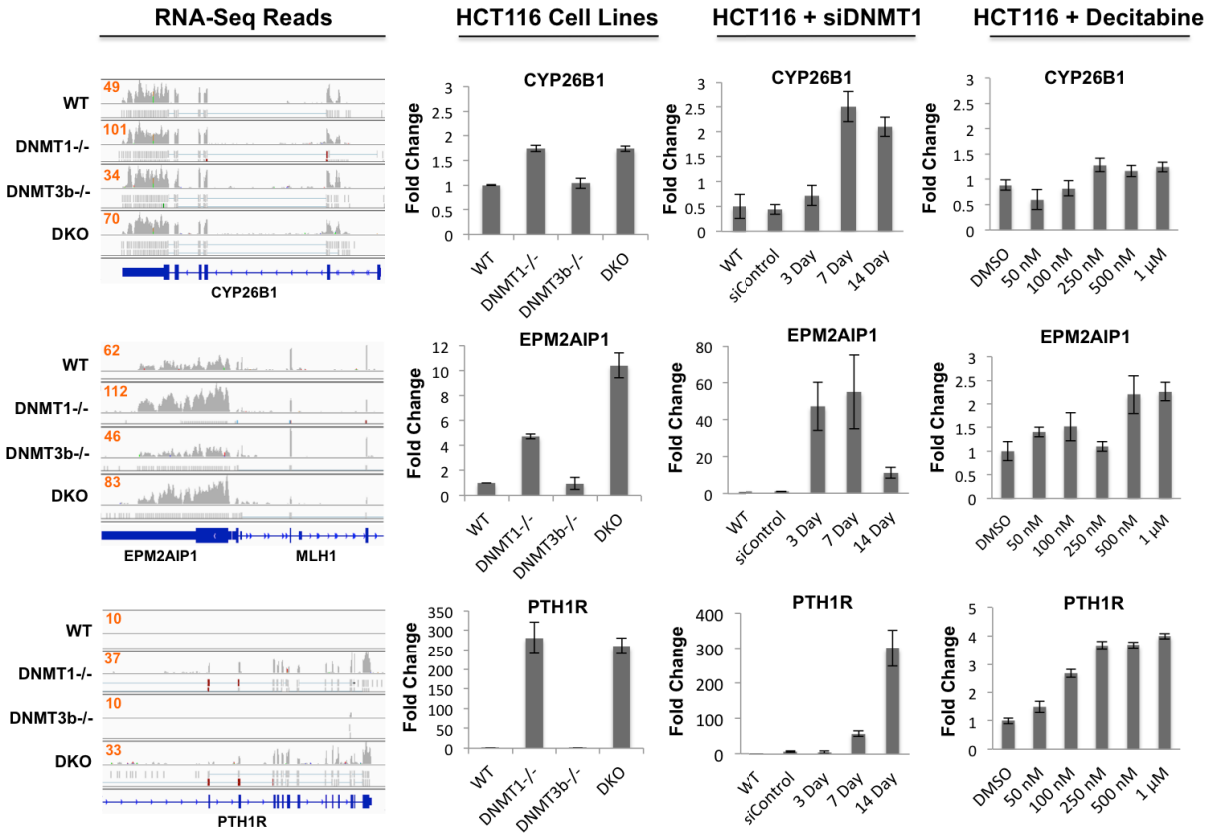


Figure 2.16: ***DNMT1 Requiring Methylation Mediates Gene Repression - Additional Examples.*** Quantitative RT-PCR confirms DNMT1 mediated gene repression in CYP26B1, EPM2AIP1, and PTH1R. Transient silencing of DNMT1 through RNAi and Decitabine recovers gene expression.



2.4 DISCUSSION AND CONCLUSIONS

In this chapter we examined the methylation profile of a set of HCT116 isogenic cell lines. We found that DNMT1 was mostly responsible for promoter hypermethylation, while DNMT3b primarily methylated CpG islands in the gene body. This fact alone points to a distinct preference for DNMT1 and DNMT3b in distinct parts of the genome.

We also observed numerous genomic regions that showed hypomethylation in HCT116-DKO alone. These regions constitute the vast majority of our hits. We postulate that in these regions, DNMT1 and DNMT3b have similar mechanisms of action, acting redundantly to maintain methylation.

Results from MBD-seq revealed that DNMT1 is most clearly responsible for differential methylation in the majority of regions where methylation depends on one methyltransferase. These DNMT1 requiring regions were mostly concentrated in promoter CpG islands, suggesting that DNMT1 mediated methylation is directly involved in gene silencing. DNMT1 was also shown to localize to regions of differential methylation, confirming its direct role in methylating these regions. Bisulfite sequencing of HCT116 transfected with siDNMT1 showed demethylation in clones corresponding to DNMT1 methylation dependent regions only, indicating that our observations are independent of clonal selection.

DNMT1's role in gene repression was confirmed when we compared our MBD-seq results to reads generated from RNA-seq of the four cell lines. For a subset of 14 genes,

we discovered that loss of methylation through loss of DNMT1 alone resulted in gene expression. When siDMT1 and Decitabine were used to transiently lower DNMT1 expression, we saw a corresponding increase in gene expression. These findings point to a direct role for DNMT1 in gene silencing through promoter CpG hypermethylation.

Interestingly, regions dependent on DNMT3b for methylation were much less prominent, and when they are present, are weaker hits than their DNMT1 dependent counterparts. This could suggest that while DNMT1 can mediate gene expression independent of other DNMTs, DNMT3b mostly works in conjunction with other methyltransferases. With the exception of the HCT116-DKO, which has a small residual amount of DNMT3a expression, all the cell lines overexpress DNMT3a, a protein that is already known to work in conjunction with DNMT3b as a de novo methyltransferase. Perhaps in certain genomic regions DNMT3a acts like a 'backup' methyltransferase in the absence of DNMT3b. Further research is needed to tease out these mechanisms.

Examining our genome wide sequencing results as a whole, it becomes apparent that each DNMT does not have the same function in all genomic regions. The interplay between the proteins seems complex. While we have successfully elucidated a mechanism by which DNMT1 mediates gene repression, we also uncovered over 10,000 additional regions showing hypomethylation in HCT116-DKO alone. DNMT3a may play a greater role in these regions as well, covering for the function of DNMT1 or DNMT3b when they are absent. In the future, a closer examination through ChIP-seq may help determine the genomic localization of all the active mammalian methyltransferases.

III. DNMT1 MEDIATES GENE EXPRESSION INDEPENDENT OF DNA METHYLATION

3.1 Introduction

In the preceding chapter, we discussed the role of DNMT1 in gene repression by methylating promoter CpG sites. However, it is possible that DNMT1 uses non-catalytic mechanisms of activity to modulate gene expression. The existence of functional roles for the N-terminal region of DNMT1 has been hinted at in the literature. For example, depletion of DNMT1 by antisense knockdown in A549 and HEK293 cells induces gene expression in a DNA methylation-independent mechanism [88]. DNMT1 was also found to activate Sp1 response elements independent of increased Sp1 expression, indicating that DNMT1 must be acting directly or indirectly on proteins interacting with Sp1 and modulating their trans-activation activity [88]. In a separate report, the N-terminal region of DNMT1 was also found to mediate E-Cadherin through its direct interaction with the E-cadherin transcriptional repressor SNAIL1 [89].

DNMT1 has been well characterized as a maintenance methyltransferase that can mediate gene expression primarily through DNA methylation. However, evidence also suggests that DNMT1 mediated gene repression may not always be dependent on methylation. Treatment of some cells by a DNMT1 inhibitor 5-aza-2′-deoxycytidine (Decitabine), causes increases in expression of some genes without evident changes in DNA methylation [90, 91]. Since Decitabine treatment is also known to cause DNMT1

degradation [92, 93], the repressive ability of the protein at specific silenced and unmethylated genes may be due to a non-catalytic function of DNMT1 potentially associated with additional proteins, or through indirect secondary effects mediated by genes that are primarily demetylated and activiated by Decitabine. Additionally, DNMT1 has been shown to modulate gene activity independent of its catalytic domain [94]. When a catalytically inactive DNMT1 was expressed in HCT116-*DNMT1-/-* cells, certain target genes were still repressed [95]. In the same publication, the histone demethylase LSD1 was recruited by DNMT1 to specific promoters and partially mediated expression of these genes.

In this chapter, we will identify genes for which DNMT1 acts as a transcriptional repressor independent of DNA methylation, and examine some of the complexities and additional work needed to more definitively test the hypotheses raised by the work presented here.

3.2 Methods

3.2.1 Transient expression of DNMT1 full length and mutant constructs

The following DNMT1 full length and mutant constructs were previously generated by a member of our lab (Gangadharan S, unpublished): DNMT1 full length, C1226S, C670S, K1535A, C1001S, C664S, and C686S, where C----S denotes a cysteine to serine mutation at amino acid ----. 2µgs of DNMT1 full length or mutant construct was transfected into HCT116-*DNMT1-/-* cells with 6µL Lipofectamine 2000 (Invitrogen) resuspended in 100µL of Optimem (Life Technologies). Cells were plated at 60% confluency in 6-well tissue culture plates 24 hours prior to transfection. Growth media (McCoy's + 10% FBS) was replaced with McCoy's serum free media immediately prior to transfection. Lipid and DNA reagents were allowed to complex in Optimem at room temperature for 15 minutes before being added to cells. Five hours post transfection, complexes were removed and media changed to McCoy's with 10% FBS. Cells were harvested for analysis 48 hours post transfection.

3.3 Results

3.3.1 DNMT1 can influence gene expression independent of methylation

As described in Chapter II of this dissertation, RNA-seq was performed on HCT116-WT, HCT116-*DNMT1-/-*, HCT116-*DNMT3b-/-*, and HCT116-DKO and compared to the differentially methylated regions identified by MBD-seq (Table 3.1). Our analysis identified a subset of genes that were activiated in DNMT1-/- compared to WT cells, but that did not show any change in methylation (either both methylated, or both unmethylated). For instance, while MAGEA4 showed absence of promoter methylation across all 4 HCT116 cell lines, disruption of DNMT1 in MAGEA4 precipitates an increase in MAGEA4 expression. These results were validated by quantitative RT-PCR. Furthermore, upon transient depletion of DNMT1 via RNAi or Decitabine, MAGEA4 expression was restored, indicating that DNMT1 acts as a gene repressor independent of methylation activity (figure 3.1). In total, we confirmed 10 genes that displayed similar patterns of DNMT1 mediated, methylation independent repression (figure 3.2 for additional examples).

Interestingly, we also identified two genes, TBX3 and FLNC, which were equally methylated in both the WT and DNMT-/- cells, but became transcriptionally activated in the absence of DNMT1 (figure 3.3 and 3.4). Quantitative RT-PCR confirmed the RNA-seq results for both genes, and RNAi and Decitabine mediated depletion of DNMT1 led to expression of TBX3 and FLNC.

From the data set shown above, we hypothesized that DNMT1 may play a role in the direct repression of these genes independent of it catalytic activity. However, the mechanisms by which DNMT1 can mediate repression remains to be determined.

Closer examination of DNMT1's methylation independent role in gene control requires knowledge of its regulatory and catalytic domains. It is well known that that specificity of DNMT1 for methylation sites resides in its catalytic domain [94]. There is also evidence that suggests an autoinhibitory role for the RFTS domain in DNMT1 [96]. Additionally, there is emerging evidence that the CXXC domain of DNMT1 is essential for its enzymatic activity [97]. Taking these hints from literature, we developed a series of DNMT1 mutation constructs with the objective of asking which one of these constructs, when overexpressed in HCT116-*DNMT1-/-* cells, can rescue gene repression of the genes described above (Gangadharan S, unpublished). Failure to rescue would indicate that the particular mutated amino acid or deleted domain is important for mediated trans repression independent of methylation. A series of 9 DNMT1 mutant constructs were generated; DNMT1 full length, CXXC, C1226S (cysteine to serine mutation at amino acid 1226), C670S, K1535A, C1001S, C664S, and C686S (Figure 3.5). All mutant constructs were successfully transfected and overexpressed in HCT116-*DNMT1-/-* cells.

Initially, when the mutant constructs were introduced into HCT116-*DNMT1-/-*, the results were promising. The full length DNMT1 as well as catalytically dead constructs repressed MAGEA4, while the vector control (TAP vector) behaved similarly to the HCT116-*DNMT1-/-* cells. There was even a slight increase in MAGEA4 expression with

the C670S mutant construct, suggesting involvement of this cysteine in repression. However, subsequent experiments were plagued by variability. Data from 3 separate trials of 5 replicates each for the MAGEA4 gene are shown in (Figure 3.6). Not only did we observe variability between experiments, the standard error within each experiment covered a very broad range. We suspected that the variability could be attributed to differences in the levels of expression achieved in each experiment, and from this, we concluded that a different approach utilizing stable cell lines containing DNMT1 deletion mutants is necessary, and can be pursued in future studies.

Table 3.1: ***DNMT1 Influences Gene Expression Independent of Methylation.*** Two scenarios exist for methylation independent, DNMT1 mediated repression. (A) Methylation is present in both WT and HCT116-*DNMT1-/-,* and the gene is induced in HCT116-*DNMT1-/-* compared to WT, or (B) Absence of methylation in both HCT116-WT and HCT116-DNMT1-/-, with induction of expression in HCT116-*DNMT1-/-*

A.

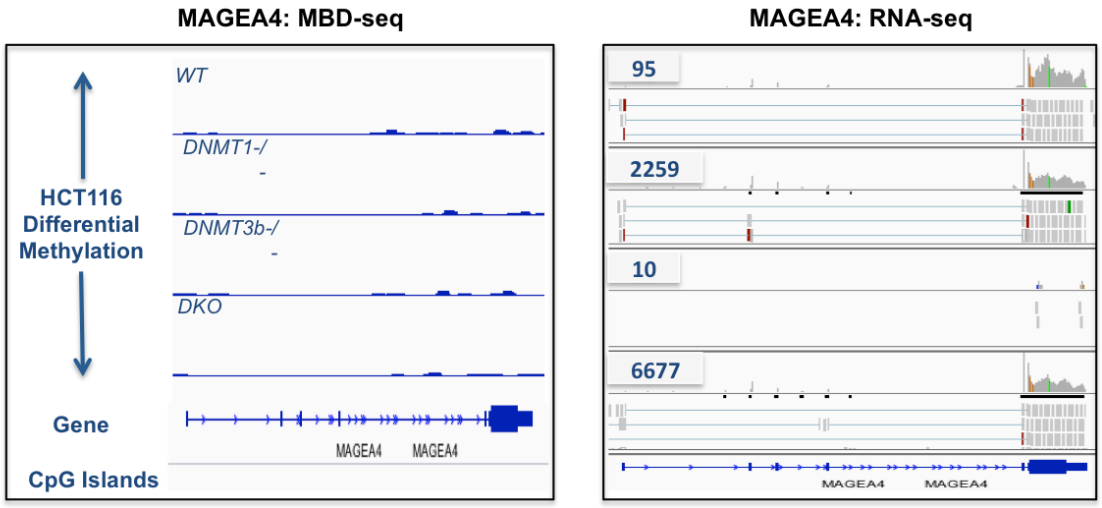
Gene	Methylation				Expression							
	(+) present, (−) absent				(+) present, (−) absent				(Number of Reads)			
	wt	1-/-	3b-/-	dko	wt	1-/-	3b-/-	dko	wt	1-/-	3b-/-	dko
TBX3	+	+	+	−	−	+	−	+	36	264	31	13
FLNC	+	+	+	−	−	+	−	+	10	215	10	423

B.

TXNIP	−	−	−	−	−	+	−	+	39	1601	10	235
PEG10	−	−	−	−	−	+	−	+	22	1191	12	239
FN1	−	−	−	−	−	+	−	+	35	393	26	22
HEY1	−	−	−	−	−	+	−	+	27	65	29	14
CTCFL	−	−	−	−	−	+	−	+	10	207	10	443
MAGEA4	−	−	−	−	−	+	−	+	95	2259	10	6677
ZNF704	−	−	−	−	−	+	−	+	14	121	39	32
NXT2	−	−	−	−	−	+	−	+	119	823	97	702
NOTUM	−	−	−	−	−	+	−	+	110	606	95	243
HCLS1	−	−	−	−	−	+	−	+	99	390	80	314

Figure 3.1: ***DNMT1 Mediates MAGEA4 Repression in the Absence of Methylation.*** (A) DNMT1 mediates methylation independent repression of MAGEA4. MBD-seq shows lack of MAGEA4 methylation in all HCT116 cell lines, but RNA-seq results reveal MAGEA4 induction only in the absence of DNMT1. (B) Quantitative RT-PCR confirms MAGEA4 expression in the absence of DNMT1. Transient depletion of DNMT1 through RNAi and Decitabine recovers MAGEA4 expression.

A.



B.

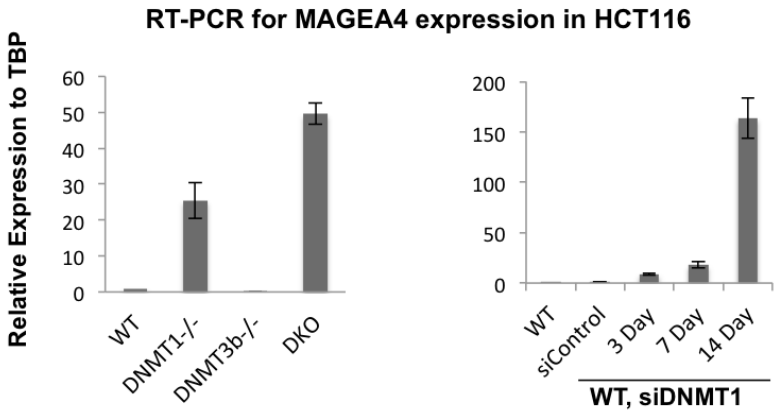


Figure 3.2: ***DNMT1 Mediates Gene Repression in the Absence of Methylation – Additional Examples.*** DNMT1 mediates methylation independent repression of PEG10, HEY1, and NOTUM. Transient depletion of DNMT1 through RNAi and Decitabine recovers gene expression.

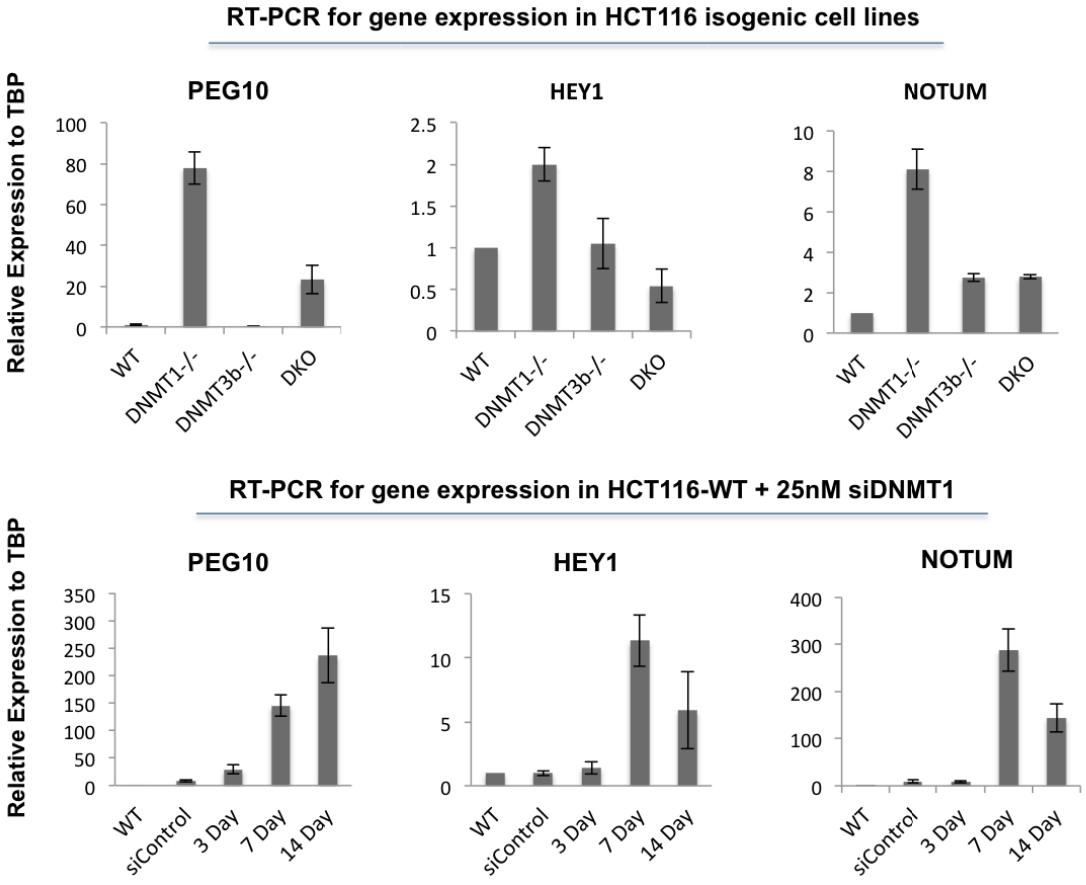
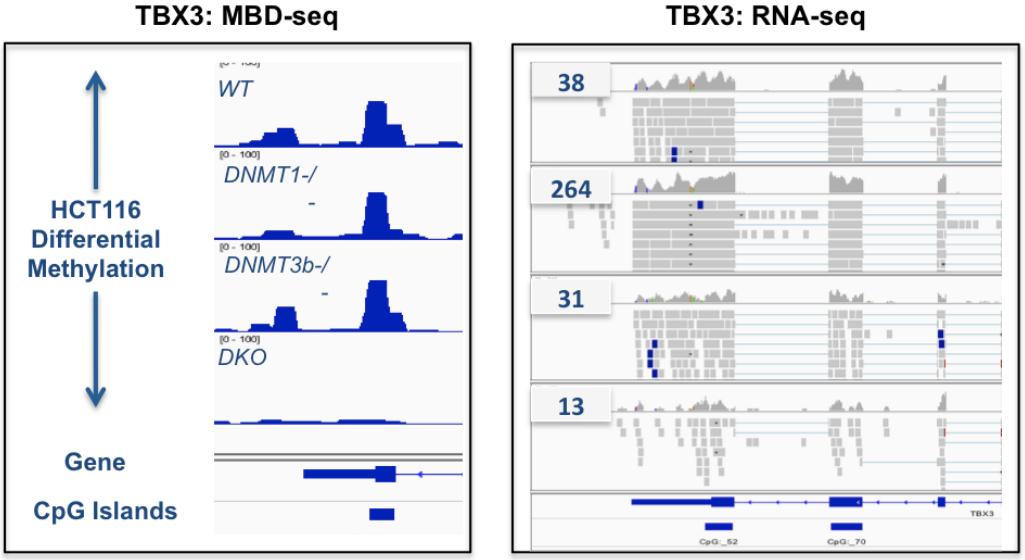


Figure 3.3: ***DNMT1 Mediates TBX3 Expression Independent of Methylation.*** (A) Although TBX3 shows hypermethylation in every HCT116 cell line except DKO, RNA-seq reveals TBX3 expression in HCT116-*DNMT1-/-* alone. (B) Quantitative RT-PCR confirms DNMT1 mediated gene repression. Transient silencing of DNMT1 through RNAi and Decitabine recovers TBX3 expression.

A.



B.

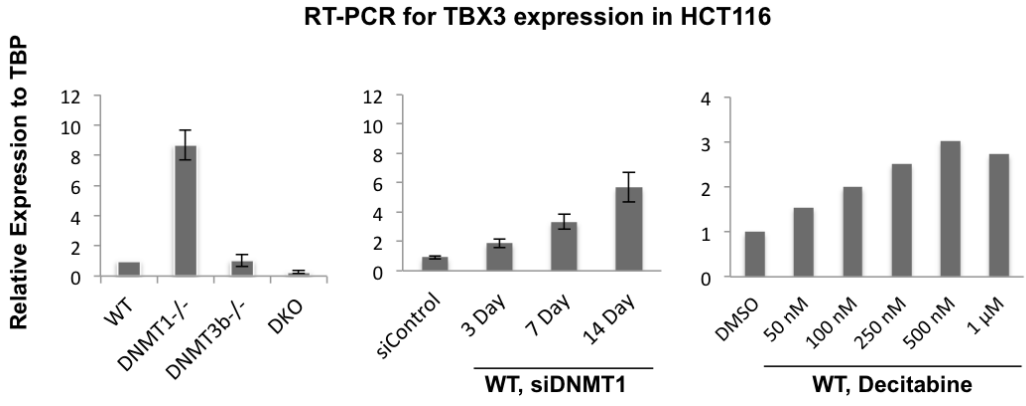


Figure 3.4: ***DNMT1 Mediates FLNC Expression Independent of Methylation.*** (A) Although FLNC shows hypermethylation in every HCT116 cell line except HCT116-DKO, RNA-seq reveals FLNC expression only in absence of DNMT1. (B) Quantitative RT-PCR confirms DNMT1 mediated gene repression. Transient silencing of DNMT1 through RNAi and Decitabine recovers FLNC expression.

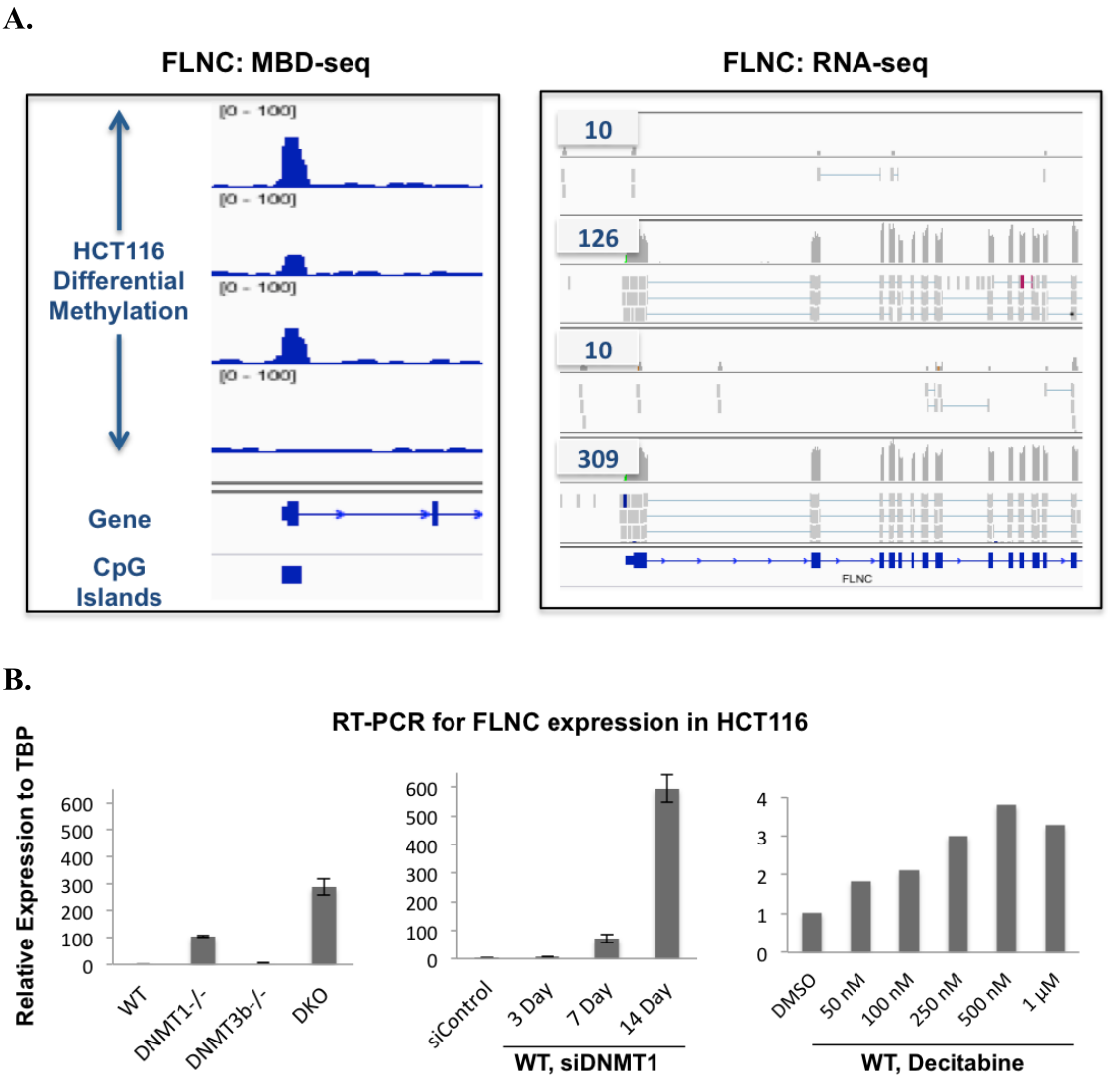
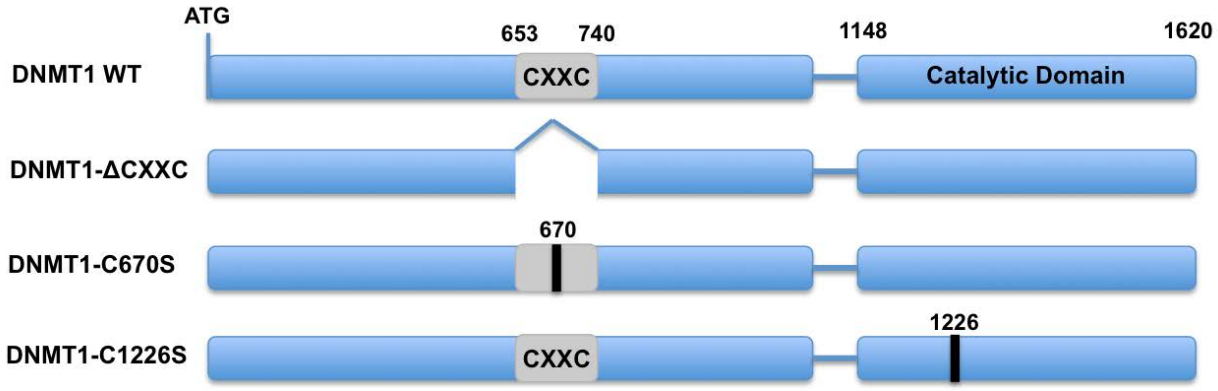


Figure 3.5: ***DNMT1 Full Length and Mutant Constructs.*** (A) Schematic of DNMT1 full length and mutant constructs. C---S denotes a cysteine to serine mutation at amino acid ---. (B) Western blot for DNMT1. Constructs were successfully overexpressed into HCT116-DNMT1-/- cells.

A.



B.

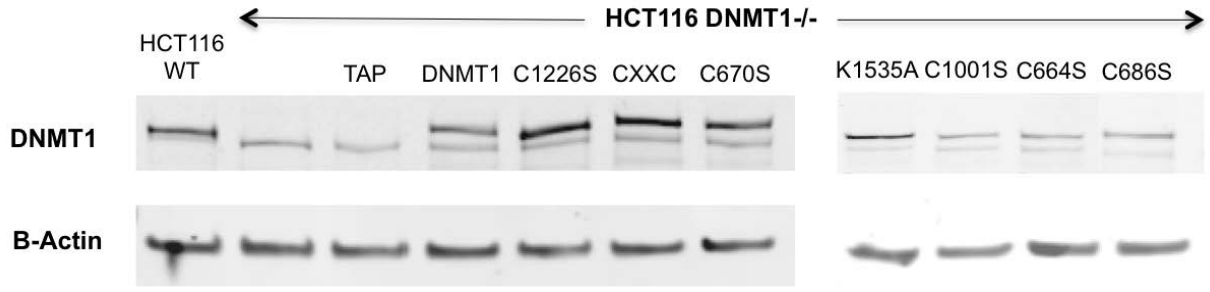
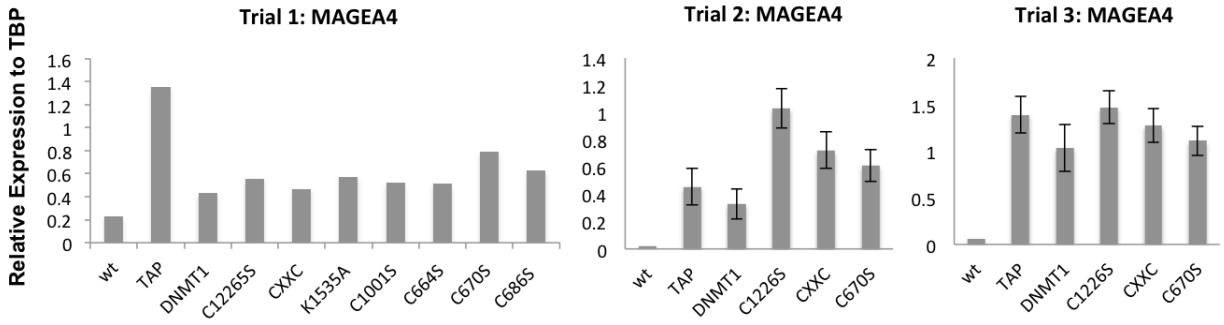


Figure 3.6: ***Changes in MAGEA4 Expression Upon Insertion of DNMT1 Full Length and Mutant Constructs.*** Successful over-expression of mutant constructs did not translate to reproducible results. Quantitative RT-PCR revealed variable expression levels between each trial. Error bars in Trials 2 and 3 represent the standard error from 5 replicates.



3.4 Discussion and Conclusions

Our genome wide profiles of methylation and gene expression reveal a set of genes repressed by DNMT1 independent of DNA methylation. We found a subset of genes that were unmethylated in all cell lines, but selectively expressed upon DNMT1 disrupted cells. Transient depletion of DNMT1 through RNAi and Decitabine independently confirmed these results.

The suggestion that DNMT1 has repressive functions independent of promoter hypermethylation points to a multifaceted role for DNMT1 in gene regulation. In order to understand the full role of this protein in gene expression control, we will have to examine how DNMT1 combines its transcriptional repression mechanism with DNA methyltransferase activity in gene silencing. This may be more complicated than it appears, as we have shown that DNMT1 mediates gene repression in varying ways in different regions of the genome. A number of studies have shown that DNMT1 interacts directly with HDAC1 and HDAC2, indicating a role for DNMT1 as part of a repressive complex [65, 66, 68]. Furthermore, a recent study of tumor suppressor gene RASSF1A revealed that HBx promoted recruitment of DNMT1 to RASSF1A promoter regions, which occured independent of RASSF1A promoter methylation [98]. Therefore, one possible way DNMT1 may mediate methylation independent gene repression is through recruitment of other histone modifying proteins.

Initially, we hoped to identify the portion of DNMT1 responsible for methylation independent of gene repression by transiently transfecting a set of DNMT1 full length and mutant constructs into HCT116-*DNMT1-/-* cells. However, it became apparent that transient incorporation of mutant constructs produced variable and unreliable results. A likely explanation is that the short time frame for disruption is not sufficient for analysis of downstream effects. Inherent variations between experiments are also a possibility. Because each experiment involves numerous steps, anything from transfection reagent toxicity to variable RNA to cDNA conversion rates can cause significant variability in the final measures. Hence, it may be worthwhile to create a set of stable cell lines containing various truncated or mutated versions of DNMT1 knocked in to the endogenous locus, or inserted randomly in the genome. These cell lines may help identify which portion(s) of DNMT1 are responsible for methylation independent repression.

Identifying DNMT1's methylation independent role in gene control has translational implications. For instance, a number of cancer therapeutics have been developed to reverse abnormal gene silencing associated with hypermethylation in promoter CpG islands [99, 100]. DNA demethylating drugs such as 5-azacytidine and Decitabine (5-aza-2′deoxycytidine) are popular drugs that block the catalytic function of DNMT1. However, inhibiting DNA methylation has also been shown to induce genes that promote metastasis [101-103]. Additionally, hypomethylation has also been shown to promote T cell lymphoma in mice hypomorphic for DNMT1 [45]. Therefore, if we can elucidate the mechanism by which DNMT1 mediates methylation independent gene repression, we can unlock new targets for chemotherapeutics.

Another potential improvement for chemotherapeutics can also occur through study of DNMT1's gene repressive mechanisms. Treatment with Decitabine (5-aza-2′deoxycytidine) and 5-azacytidine may also cause changes in gene expression that spans beyond genes regulated by DNA methylation alone. Further experiments that explore this possibility will enable us to more clearly understand DNMT1's role in tumorigenesis and disease.

IV. CONCLUSION

Identifying DNMTs' methylation dependent and independent roles in gene silencing is an important aspect of epigenetic research. Historically, DNMTs have been widely implicated in methylation mediated gene regulation. As a result of extensive work in this field, various roles have been canonically assigned to each methyltransferase. While DNMT1 labeled as a maintenance methyltransferase, the DNMT3 family was implicated in de novo methylation. Our work helps to clarify the roles of the DNMT1 and DNMT3b methyltransferases in maintaining genome methylation patterns. We find that DNMT1 is required for promoter methylation and repression in a significant subset of genes and that DNMT3b is required for maintaining methylation at a number of coding exons. However, methylation of genomic regions by DNMTs appears mostly a redundant/cooperative effort- as the predominance of regions can lose their methylation only when both DNMT1 and DNMT3b are disrupted.

The presence of genomic regions methylated in HCT116-DKO provides additional areas worthy of exploration. We observed a small subset of regions that were methylated in the absence of DNMT1 and DNMT3b. These results were not unexpected, as methylation profiles of the same clone of HCT116-DKO1 cells published earlier revealed a subset of hundreds of regions that still retained a high level of DNA methylation in these cells [104]. Since DNMT3a is still present in small amounts in HCT116-DKO, its role in these methylated sites may be worthy of exploration. Additionally, the HCT116-DNMT1 and HCT116-DKO cells used in the current and prior studies continue to retain and express a

hypomorphic isoform of DNMT1 with extremely reduced, yet measurable, catalytic activity [107]. Thus the residual methylation in the DKO cells may be partially maintained by this hypomorphic residual activity.

Of all the DNMTs, DNMT1 appears most likely to play a major role in cancer development. The enzyme has been reported to be required for intestinal polyp development in $Apc^{Min/+}$ mice and for tobacco carcinogen-induced murine lung cancer development *in vivo* [62, 105]. Low levels of DNMT1 can also result in cancer development. Mice carrying one disrupted DNMT1 allele and one hypomorphic DNMT1 allele only retain 10% of normal DNMT activity, and are reported to exhibit genomic instability and to develop T-cell lymphomas [44, 45]. We have provided evidence that DNMT1 mediated promoter methylation appears solely responsible for gene repression in a number of genes. If aberrant DNMT1 activity is indeed an important precursor to carcinogenesis, our results indicate that DNMT1 methylation mediated gene silencing is one way the enzyme may contribute to cancer development. The interesting observation that specific regions of the genome require DNMT1 for maintenance of methylation suggests that such regions might be particularly prone for DNMT inhibitor based gene reactivation. Such knowledge may be helpful for developing strategies for pharmacodynamic tracking of DNMT inhibitor responses.

It is only recently that DNMT1's N-terminal activity was revealed in literature. In our studies, DNMT1's methylation independent gene repression activity was discovered for several new genes. Identifying the portion(s) of DNMT1 responsible for gene repression

in this manner would open up new possibilities for targeting DNMT1. Although demethylating drugs such as Azacitidine and Decitabine have existed for numerous years, their use involves hematologic side effects that have limited the ability to create effective dosing regimens. Therefore, we hope our data paints a clearer portrait of DNMT1 activity, leading to identification of new insights into cancer pathobiology and therapeutic strategies.

IV. References

1. Wolffe, A.P. and M.A. Matzke, *Epigenetics: regulation through repression.* Science, 1999. **286**(5439): p. 481-6.
2. Wilmut, I., *Cloning for medicine.* Sci Am, 1998. **279**(6): p. 58-63.
3. Jones, P.A. and M.L. Gonzalgo, *Altered DNA methylation and genome instability: a new pathway to cancer?* Proc Natl Acad Sci U S A, 1997. **94**(6): p. 2103-5.
4. Tilghman, S.M., *The sins of the fathers and mothers: genomic imprinting in mammalian development.* Cell, 1999. **96**(2): p. 185-93.
5. Li, E., T.H. Bestor, and R. Jaenisch, *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.* Cell, 1992. **69**(6): p. 915-26.
6. Tate, P., W. Skarnes, and A. Bird, *The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse.* Nat Genet, 1996. **12**(2): p. 205-8.
7. Chang, S.C., et al., *Mechanisms of X-chromosome inactivation.* Front Biosci, 2006. **11**: p. 852-66.
8. Delaval, K. and R. Feil, *Epigenetic regulation of mammalian genomic imprinting.* Curr Opin Genet Dev, 2004. **14**(2): p. 188-95.
9. Heard, E., *Recent advances in X-chromosome inactivation.* Curr Opin Cell Biol, 2004. **16**(3): p. 247-55.
10. Hore, T.A., R.W. Rapkins, and J.A. Graves, *Construction and evolution of imprinted loci in mammals.* Trends Genet, 2007. **23**(9): p. 440-8.

11. Yen, Z.C., et al., *A cross-species comparison of X-chromosome inactivation in Eutheria.* Genomics, 2007. **90**(4): p. 453-63.

12. Straub, T. and P.B. Becker, *Dosage compensation: the beginning and end of generalization.* Nat Rev Genet, 2007. **8**(1): p. 47-57.

13. Sha, K., *A mechanistic view of genomic imprinting.* Annu Rev Genomics Hum Genet, 2008. **9**: p. 197-216.

14. Mukherjee, S., *The emperor of all maladies : a biography of cancer*. Large print ed. 2010, Waterville, Me.: Thorndike Press.

15. Hotchkiss, R.D., *The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography.* J Biol Chem, 1948. **175**(1): p. 315-32.

16. Fitzpatrick, D.R. and C.B. Wilson, *Methylation and demethylation in the regulation of genes, cells, and responses in the immune system.* Clin Immunol, 2003. **109**(1): p. 37-45.

17. Yang, X., et al., *Nuclear reprogramming of cloned embryos and its implications for therapeutic cloning.* Nat Genet, 2007. **39**(3): p. 295-302.

18. Reik, W., *Stability and flexibility of epigenetic gene regulation in mammalian development.* Nature, 2007. **447**(7143): p. 425-32.

19. Hemberger, M., W. Dean, and W. Reik, *Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal.* Nat Rev Mol Cell Biol, 2009. **10**(8): p. 526-37.

20. Sweatt, J.D., *Experience-dependent epigenetic modifications in the central nervous system.* Biol Psychiatry, 2009. **65**(3): p. 191-7.

21. Zhang, Y., et al., *Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail.* Nucleic Acids Res, 2010. **38**(13): p. 4246-53.

22. Jones, P.A. and S.B. Baylin, *The epigenomics of cancer.* Cell, 2007. **128**(4): p. 683-92.

23. Jones, P.A. and S.B. Baylin, *The fundamental role of epigenetic events in cancer.* Nat Rev Genet, 2002. **3**(6): p. 415-28.

24. Feinberg, A.P. and B. Tycko, *The history of cancer epigenetics.* Nat Rev Cancer, 2004. **4**(2): p. 143-53.

25. Lister, R., et al., *Human DNA methylomes at base resolution show widespread epigenomic differences.* Nature, 2009. **462**(7271): p. 315-22.

26. Laurent, L., et al., *Dynamic changes in the human methylome during differentiation.* Genome Res, 2010. **20**(3): p. 320-31.

27. Takai, D. and P.A. Jones, *Origins of bidirectional promoters: computational analyses of intergenic distance in the human genome.* Mol Biol Evol, 2004. **21**(3): p. 463-7.

28. Saxonov, S., P. Berg, and D.L. Brutlag, *A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters.* Proc Natl Acad Sci U S A, 2006. **103**(5): p. 1412-7.

29. Zhang, Y., et al., *DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution.* PLoS Genet, 2009. **5**(3): p. e1000438.

30. Hodges, E., et al., *High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing.* Genome Res, 2009. **19**(9): p. 1593-605.

31. Riggs, A.D., *X inactivation, differentiation, and DNA methylation.* Cytogenet Cell Genet, 1975. **14**(1): p. 9-25.

32. Holliday, R. and J.E. Pugh, *DNA modification mechanisms and gene activity during development.* Science, 1975. **187**(4173): p. 226-32.

33. Okano, M., S. Xie, and E. Li, *Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases.* Nat Genet, 1998. **19**(3): p. 219-20.

34. Gowher, H. and A. Jeltsch, *Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites.* J Mol Biol, 2001. **309**(5): p. 1201-8.

35. Goll, M.G., et al., *Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2.* Science, 2006. **311**(5759): p. 395-8.

36. Jurkowska, R.Z., T.P. Jurkowski, and A. Jeltsch, *Structure and function of mammalian DNA methyltransferases.* Chembiochem, 2011. **12**(2): p. 206-22.

37. Miremadi, A., et al., *Cancer genetics of epigenetic genes.* Hum Mol Genet, 2007. **16 Spec No 1**: p. R28-49.

38. Fatemi, M., et al., *The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the*

enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. J Mol Biol, 2001. **309**(5): p. 1189-99.

39. Goyal, R., R. Reinhardt, and A. Jeltsch, *Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase.* Nucleic Acids Res, 2006. **34**(4): p. 1182-8.

40. Leonhardt, H., et al., *A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei.* Cell, 1992. **71**(5): p. 865-73.

41. Hermann, A., R. Goyal, and A. Jeltsch, *The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites.* J Biol Chem, 2004. **279**(46): p. 48350-9.

42. Robertson, K.D., *DNA methylation, methyltransferases, and cancer.* Oncogene, 2001. **20**(24): p. 3139-55.

43. Chen, T., et al., *Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells.* Nat Genet, 2007. **39**(3): p. 391-6.

44. Eden, A., et al., *Chromosomal instability and tumors promoted by DNA hypomethylation.* Science, 2003. **300**(5618): p. 455.

45. Gaudet, F., et al., *Induction of tumors in mice by genomic hypomethylation.* Science, 2003. **300**(5618): p. 489-92.

46. Okano, M., et al., *DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development.* Cell, 1999. **99**(3): p. 247-57.

47. Hansen, R.S., et al., *The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome.* Proc Natl Acad Sci U S A, 1999. **96**(25): p. 14412-7.

48. Kareta, M.S., et al., *Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L.* J Biol Chem, 2006. **281**(36): p. 25893-902.

49. Bourc'his, D., et al., *Dnmt3L and the establishment of maternal genomic imprints.* Science, 2001. **294**(5551): p. 2536-9.

50. Feltus, F.A., et al., *Predicting aberrant CpG island methylation.* Proc Natl Acad Sci U S A, 2003. **100**(21): p. 12253-8.

51. Jair, K.W., et al., *De novo CpG island methylation in human cancer cells.* Cancer Res, 2006. **66**(2): p. 682-92.

52. Jeong, S., et al., *Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA.* Mol Cell Biol, 2009. **29**(19): p. 5366-76.

53. Jones, P.A. and G. Liang, *Rethinking how DNA methylation patterns are maintained.* Nat Rev Genet, 2009. **10**(11): p. 805-11.

54. Holliday, R., *Epigenetic inheritance based on DNA methylation.* EXS, 1993. **64**: p. 452-68.

55. Jones, P.A. and P.W. Laird, *Cancer epigenetics comes of age.* Nat Genet, 1999. **21**(2): p. 163-7.

56. Herman, J.G., *Hypermethylation of tumor suppressor genes in cancer.* Semin Cancer Biol, 1999. **9**(5): p. 359-67.

57. Esteller, M., et al., *DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis.* Hum Mol Genet, 2001. **10**(26): p. 3001-7.

58. Merlo, A., et al., *5′ CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers.* Nat Med, 1995. **1**(7): p. 686-92.

59. Rhee, I., et al., *DNMT1 and DNMT3b cooperate to silence genes in human cancer cells.* Nature, 2002. **416**(6880): p. 552-6.

60. Rhee, I., et al., *CpG methylation is maintained in human cancer cells lacking DNMT1.* Nature, 2000. **404**(6781): p. 1003-7.

61. Laird, P.W., et al., *Suppression of intestinal neoplasia by DNA hypomethylation.* Cell, 1995. **81**(2): p. 197-205.

62. Eads, C.A., A.E. Nickel, and P.W. Laird, *Complete genetic suppression of polyp formation and reduction of CpG-island hypermethylation in Apc(Min/+) Dnmt1-hypomorphic Mice.* Cancer Res, 2002. **62**(5): p. 1296-9.

63. Bachman, K.E., M.R. Rountree, and S.B. Baylin, *Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin.* J Biol Chem, 2001. **276**(34): p. 32282-7.

64. Baylin, S.B., et al., *Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer.* Hum Mol Genet, 2001. **10**(7): p. 687-92.

65. Rountree, M.R., K.E. Bachman, and S.B. Baylin, *DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci.* Nat Genet, 2000. **25**(3): p. 269-77.

66. Robertson, K.D., et al., *DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters.* Nat Genet, 2000. **25**(3): p. 338-42.

67. Burgers, W.A., F. Fuks, and T. Kouzarides, *DNA methyltransferases get connected to chromatin.* Trends Genet, 2002. **18**(6): p. 275-7.

68. Fuks, F., et al., *DNA methyltransferase Dnmt1 associates with histone deacetylase activity.* Nat Genet, 2000. **24**(1): p. 88-91.

69. Fuks, F., et al., *Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription.* EMBO J, 2001. **20**(10): p. 2536-44.

70. De Marzo, A.M., et al., *Human prostate cancer precursors and pathobiology.* Urology, 2003. **62**(5 Suppl 1): p. 55-62.

71. De Marzo, A.M., et al., *Inflammation in prostate carcinogenesis.* Nat Rev Cancer, 2007. **7**(4): p. 256-69.

72. Sfanos, K.S. and A.M. De Marzo, *Prostate cancer and inflammation: the evidence.* Histopathology, 2012. **60**(1): p. 199-215.

73. Bastian, P.J., et al., *[GSTP1 CpG island hypermethylation as a molecular marker of prostate cancer].* Urologe A, 2004. **43**(5): p. 573-9.

74. Jacinto, F.V. and M. Esteller, *Mutator pathways unleashed by epigenetic silencing in human cancer.* Mutagenesis, 2007. **22**(4): p. 247-53.

75. Suzuki, T. and N. Miyata, *Non-hydroxamate histone deacetylase inhibitors.* Curr Med Chem, 2005. **12**(24): p. 2867-80.

76. Laird, P.W., *Principles and challenges of genomewide DNA methylation analysis.* Nat Rev Genet, 2010. **11**(3): p. 191-203.

77. Yegnasubramanian, S., et al., *Chromosome-wide mapping of DNA methylation patterns in normal and malignant prostate cells reveals pervasive methylation of*

gene-associated and conserved intergenic sequences. BMC Genomics, 2011. **12**: p. 313.

78. Jin, B., et al., *Linking DNA methyltransferases to epigenetic marks and nucleosome structure genome-wide in human tumor cells.* Cell Rep, 2012. **2**(5): p. 1411-24.

79. Jadhav, R.R., et al., *Genome-wide DNA methylation analysis reveals estrogen-mediated epigenetic repression of metallothionein-1 gene cluster in breast cancer.* Clin Epigenetics, 2015. **7**(1): p. 13.

80. Cross, S.H., et al., *Purification of CpG islands using a methylated DNA binding column.* Nat Genet, 1994. **6**(3): p. 236-44.

81. Jorgensen, H.F., et al., *Engineering a high-affinity methyl-CpG-binding protein.* Nucleic Acids Res, 2006. **34**(13): p. e96.

82. Bock, C., et al., *Quantitative comparison of genome-wide DNA methylation mapping technologies.* Nat Biotechnol, 2010. **28**(10): p. 1106-14.

83. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. **9**(9): p. R137.

84. Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov, *Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.* Brief Bioinform, 2013. **14**(2): p. 178-92.

85. Robinson, J.T., et al., *Integrative genomics viewer.* Nat Biotechnol, 2011. **29**(1): p. 24-6.

86. Ji, X., et al., *CEAS: cis-regulatory element annotation system.* Nucleic Acids Res, 2006. **34**(Web Server issue): p. W551-4.

87. Shin, H., et al., *CEAS: cis-regulatory element annotation system.* Bioinformatics, 2009. **25**(19): p. 2605-6.

88. Milutinovic, S., et al., *DNA methyltransferase 1 knock down induces gene expression by a mechanism independent of DNA methylation and histone deacetylation.* J Biol Chem, 2004. **279**(27): p. 27915-27.

89. Espada, J., et al., *Regulation of SNAIL1 and E-cadherin function by DNMT1 in a DNA methylation-independent context.* Nucleic Acids Res, 2011. **39**(21): p. 9194-205.

90. Ohm, J.E., et al., *A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing.* Nat Genet, 2007. **39**(2): p. 237-42.

91. Mohammad, H.P. and S.B. Baylin, *Linking cell signaling and the epigenetic machinery.* Nat Biotechnol, 2010. **28**(10): p. 1033-8.

92. Patel, K., et al., *Targeting of 5-aza-2'-deoxycytidine residues by chromatin-associated DNMT1 induces proteasomal degradation of the free enzyme.* Nucleic Acids Res, 2010. **38**(13): p. 4313-24.

93. Ghoshal, K., et al., *5-Aza-deoxycytidine induces selective degradation of DNA methyltransferase 1 by a proteasomal pathway that requires the KEN box, bromo-adjacent homology domain, and nuclear localization signal.* Mol Cell Biol, 2005. **25**(11): p. 4727-41.

94. Bashtrykov, P., et al., *Specificity of Dnmt1 for methylation of hemimethylated CpG sites resides in its catalytic domain.* Chem Biol, 2012. **19**(5): p. 572-8.

95. Clements, E.G., et al., *DNMT1 modulates gene expression without its catalytic activity partially through its interactions with histone-modifying enzymes.* Nucleic Acids Res, 2012. **40**(10): p. 4334-46.

96. Bashtrykov, P., et al., *Targeted mutagenesis results in an activation of DNA methyltransferase 1 and confirms an autoinhibitory role of its RFTS domain.* Chembiochem, 2014. **15**(5): p. 743-8.

97. Pradhan, M., et al., *CXXC domain of human DNMT1 is essential for enzymatic activity.* Biochemistry, 2008. **47**(38): p. 10000-9.

98. Qiu, X., et al., *Upregulation of DNMT1 mediated by HBx suppresses RASSF1A expression independent of DNA methylation.* Oncol Rep, 2014. **31**(1): p. 202-8.

99. Herman, J.G. and S.B. Baylin, *Gene silencing in cancer in association with promoter hypermethylation.* N Engl J Med, 2003. **349**(21): p. 2042-54.

100. Fenaux, P., *Inhibitors of DNA methylation: beyond myelodysplastic syndromes.* Nat Clin Pract Oncol, 2005. **2 Suppl 1**: p. S36-44.

101. Rosty, C., et al., *Overexpression of S100A4 in pancreatic ductal adenocarcinomas is associated with poor differentiation and DNA hypomethylation.* Am J Pathol, 2002. **160**(1): p. 45-50.

102. Guo, Y., et al., *Regulation of DNA methylation in human breast cancer. Effect on the urokinase-type plasminogen activator gene production and tumor invasion.* J Biol Chem, 2002. **277**(44): p. 41571-9.

103. Gupta, A., et al., *Hypomethylation of the synuclein gamma gene CpG island promotes its aberrant expression in breast carcinoma and ovarian carcinoma.* Cancer Res, 2003. **63**(3): p. 664-73.

104. De Carvalho, D.D., et al., *DNA methylation screening identifies driver epigenetic events of cancer cell survival.* Cancer Cell, 2012. **21**(5): p. 655-67.

105. Belinsky, S.A., et al., *Inhibition of DNA methylation and histone deacetylation prevents murine lung cancer.* Cancer Res, 2003. **63**(21): p. 7089-93.

106. Sinclair, S.H., et al., *Global DNA methylation changes and differential gene expression in Anaplasma phagocytophilum-infected human neutrophils*. Clin Epigenetics, 2015. **7**(1): p.77.

107. Egger, G., et al., *Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival.* Proc Natl Acad Sci U S A, 2006. **103**(38): p. 14080-5.

Yung-Sheng Melody Tsui

E: melody.tsui@gmail.com □ **M:** 206.790.6779
Born: August 16, 1983 in Taipei, Taiwan

Education

Johns Hopkins University School of Medicine, Baltimore, MD
Ph.D, Pharmacology and Molecular Sciences
September 2008 – September 2015

University of Pennsylvania School of Engineering, Philadelphia, PA
B.S.E., Bioengineering
August 2001 – May 2005

Research Experience

Johns Hopkins University School of Medicine, Baltimore, MD
Position: Graduate Researcher (September 2009 – September 2015)
Mentors: William Nelson, M.D., Ph.D and Srinivasan Yegnasubramanian, M.D., Ph.D
Project: Distinct Roles for DNMT1 in Modulating the Cancer Epigenome

Harvard Medical School, Boston, MA
Position: Research Assistant (September 2005 – August 2008)
Mentor: Timothy Mitchison, Ph.D
Project: Designed and implemented two high-throughput human genome wide siRNA screens for enhancers and suppressors of mitotic Kinesin-5 inhibitor.

ICCB-Longwood Screening Facility, Boston, MA
Position: Microscopy Specialist (September 2005 – August 2008)
Project: Managed the high content imaging facility; Collaborated with researchers to optimize their assays on automated microscopes; Developed custom image analysis algorithms.

University of Pennsylvania, Philadelphia, PA
Position: Research Assistant (January 2004 – May 2005)
Mentor: Wei Guo, Ph.D
Project: Aided in experiments with budding yeast to determine the molecular basis for polarized exocytosis.

Publications and Presentations

Tsui YM, Yegnasubramanian S, Nelson WG, Distinct Roles for DNA methyltransferase I in Modulating the Cancer Epigenome, *Manuscript in preparation*

Tsui M, Haffner M, Aryee M, Nelson WG, Yegnasubramanian S. Requirement of Specific DNA Methyltransferases for Maintenance of Methylation at Specific Genomic Regions. *Multi-institutional Prostate Cancer Meeting*, Ft. Lauderdale, FL, March 2012. Poster and Talk.

Tsui M, Haffner M, Aryee M, Nelson WG, Yegnasubramanian S. Requirement of Specific DNA Methyltransferases for Maintenance of Methylation at Specific Genomic Regions. *Multi-institutional Prostate Cancer Meeting*, Ft. Lauderdale, FL, March 2011. Poster.

Tsui M. Xie T. Orth JD. Carpenter AE. Rudnicki S. Kim S. Shamu CE. Mitchison TJ. An Intermittent Live Cell Imaging Screen for siRNA Enhancers and Suppressors of a Kinesin-5 Inhibitor. PLoS One. October 2009.

Krishnan MN, Ng A, Sukumaran B, Gilfoy FD, Uchil PD, Sultana H, Brass AL, Adametz R, **Tsui M**, Qian F, Montgomery RR, Lev S, Mason PW, Koski RA, Elledge SJ, Xavier RJ, Agaisse H, Fikrig E. RNA Interference Screen for Human Genes Associated with West Nile Virus Infection. Nature. August 2008.